Essays on Institutional Investors in Corporate Bond Markets

Minchen Zheng

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

ABSTRACT

Essays on Institutional Investors in Corporate Bond Markets

Minchen Zheng

This dissertation focuses on institutional investors in corporate bond markets and their impacts on the underlying corporate bonds. The dissertation is composed of three chapters. The first chapter studies how information networks of corporate bond mutual funds may be constructed. It highlights how information flow between corporate bond mutual funds affects fund performance, herding behaviors, and the underlying corporate bond market. By examining the trading behavior of corporate bond mutual funds, I show that bond funds in more central positions of a trading network have an informational advantage that results in 0.33% higher future fund risk-adjusted return. This positive relationship becomes stronger during periods of high market uncertainty and for bond funds with more liquid assets as they can respond to the information signal with a lower asset reallocation cost. I further show that manager turnover, ranking pressure, and fund flow fluctuation drive within-fund time varying changes in network centrality.

The second chapter exploits the influence of information networks of corporate bond mutual funds on their underlying corporate bonds. I show that corporate bonds owned by highly network-central bond funds increase underlying liquidity, leading to a 3.5% decrease in future bid/ask spreads and a lower Amihud illiquidity measure. This is hypothesized to occur due to increased information efficiency, which allows for more bond specific information to be reflected in the price, and intensified herding behavior. Moreover, I construct a network to represent herding behavior by following the same

trades across quarters. A 0.3% decrease in risk-adjusted fund return is found for bond funds that have the strongest herding behavior.

The third chapter studies how corporate bond exchange-traded funds (ETF) impact the underlying corporate bond return comovement and how it relates to trading and arbitrage activities of corporate bond ETFs. The literature is silence about the effect of corporate bond ETFs on the comovement of underlying bond securities. This chapter aims to fill this gap by providing the first empirical evidence of bond return comovement driven by bond ETFs ownership. I find that bond ownership by corporate bond ETFs leads to higher bond return comovement, an increase of 0.26 in the beta of corporate bond return with respect to the aggregate bond portfolio. In contrast, bond ownership by other traditional institutional investors in the corporate bond market like bond mutual funds and insurance companies do not contribute to corporate bond return comovement. Furthermore, this chapter highlights that return comovement is driven by corporate bond ETFs' creation and redemption activities.

[this page intentionally left blank]

Contents

List of Charts, Graphs, and Illustrations

Acknowledgements

I am deeply indebted to my dissertation committee: Wei Jiang (sponsor), Suresh Sundaresan (Chair), Simona Abis, Kairong Xiao, and Kunal Sachdeva. I have been honored to have them as my advisors. I thank them for invaluable support and insightful guidance in making this dissertation a reality.

This work is dedicated to my parents for their love, endless support and prayers.

[this page intentionally left blank]

Chapter 1 : Information Networks of Corporate Bond Mutual Funds

1.1 Introduction

The size of corporate bond mutual funds has grown unprecedentedly in the last decade, tripling from 0.68 trillion in 2004 to 1.8 trillion dollars in 2018 (Investment Company Institute 2019). Holding more than 20% of total outstanding bonds, corporate bond mutual funds are one of the most active players in the corporate bond market compared to other participants who are primarily buy-and-hold investors (like insurance funds and pension funds). Corporate bond markets used to rely on intermediary dealers to provide liquidity (Harris (2015)). However, dealers retreated from market-making activities after the recent financial crisis. Since then, corporate bond mutual funds have assumed an increasingly important intermediary role in the corporate bond market Adrian, Flemming, Shachar, Vogt (2016) and Choi and Huh (2017)). Therefore, it is important to understand the trading behavior of corporate bond mutual funds and its impact on the corporate bond market. What drives their different trades? How does their trading influence fund performance? Does trading by corporate bond funds improve underlying corporate bond liquidity and increase the information embedded in prices?

Information network analysis enables us to analyze investors' heterogeneous trading behaviors effectively, as funds exist within networks that inform and constrain their behavior. Investors make their trading decision according to information they obtain through social connections and their peers (Choi and Huh (2017), Ivković and Weisbenner (2007) and Cohen, Frazzini, and Malloy (2008)). Hence, investors with access to the same sources of information through social relationships can be seen as conceptually forming an information network. Established network theory shows that investors positioned more centrally in an information network have a greater information advantage resulting from more connections and earlier access to information. (Jackson (2009), Ozsoylev and Walden (2011) and Han and Yang (2013)). Starting with Pareek (2012) and Ozsoylev, Walden, Yavuz and Bildik (2014), empirical studies have examined the investor information network in stock markets. However, little[1] is known about the information networks of bond mutual funds in corporate bond markets. In this study, I present an analysis of the corporate bond mutual funds' trading information network and its influence on fund performance and underlying corporate bond liquidity.

I construct a weighted information network of corporate bond mutual funds based on their quarterly trading behaviors from eMaxx database. Corporate bond mutual funds who trade on same corporate bond with the same direction (buy or sell) within the same quarter are purported to share a similar view or access to the same information signal, either private or public, on the specific corporate bond. This is the basis of the information linkage in the network with nodes representing each bond mutual fund. For each pairwise information edge between funds in the network, I assign the weight based

[1] Two exceptions are Falato, Hortacsu, Li and Shin (2016) and Booth, Gurun, and Zhang (2014). However, I have very different focuses. The former studies fund flow driven by asset class overlap. The latter studies the advantage of globally connected vs local financial institutions in Turkish government bond market. I focus on information network based on heterogeneous trading behaviors of corporate bond mutual fund and its impact on bond fund performance and broader corporate bond market.

on pairwise overlapping trading intensity, measured by proportions of the same trades shared by both funds each quarter. My weighted information network of corporate bond mutual funds has following two advantages.

First, edge links in my network not only captures the fact that pairwise bond funds have the same trades, but the intensity of their trade overlap with respect to their quarterly total trading volume. This model improves upon the unweighted trading networks constructed by Ozsoylev, Walden, Yavuz and Bildik (2014) and Sung (2016). Their networks cannot differentiate edge links with different trade volumes or overlapping intensities. This could disproportionally emphasize edges with thin trades and result in more noise within the graph.

Second, by using trading behavior instead of common holding positions of bond mutual funds to build the network (Pareek (2012)), the probability of capturing stale information caused by the inability to liquidate bond positions is greatly reduced. For example, bond mutual funds A and B may have different views on a bond they both hold. Fund A wants to sell the bond but fails to do so due to high transaction costs during the current quarter, but instead sells the bond in the following quarter. The network built on holding linkages still links both funds by this hard-to-sell bond and records the stale information. In contrast, the information network based on trading overcomes this issue and shows no common information linkage between A and B during the current quarter.

I use eigenvector centrality to measure each mutual fund's centrality in the information network (Bonacich (2007) and Walden (2014)). This is a superior network centrality measurement as it takes into account not only the number of connections each fund has in the network but also to which funds one connects to. The more important and connected funds that one connects to results in higher network centrality, relative to less connected bond funds. Equipped with the information network of corporate bond mutual funds, I investigate the effect of network centrality on fund performance and the underlying corporate bonds.

I first show that with informational advantages, bond funds in central positions of the trading network have higher future fund performance. This significant positive relationship between network centrality of corporate bond mutual funds and fund performance persists through various lookahead periods, ranging from one month to one year. It is consistent in various return calculations: raw return, net return after fee and fund risk adjusted return. Fund level fixed effects, total assets and trading volume control variables enable us to separate investors central in the information network from investors who merely trade much. A standard deviation increases in the eigenvector network centrality results in 0.33% increase in fund risk adjusted returns. It lends supports to theoretical predictions that an information advantage arising from high network centrality can translate into higher profitability (Walden (2014), Han and Yang (2013) and Walden (2014)).

To address the concern that same trades within quarter can be driven by style investing or benchmark index following rather than information sharing in the network, I further construct trading information network using only same trades from bond funds with different investment styles/benchmark indexes. I find that positive relationship between network centrality and fund performance remains significantly. Using the portfolio-based analysis, I further confirm this positive relationship. Bond funds sorted in the top portfolio with high network centrality subsequently outperform those with low high network centrality by 1.2% per year risk adjusted return.

Fund managers turnover, ranking pressure, and fund flow fluctuation are discovered to drive within-fund time-variant changes in network centrality. Hiring new fund managers or firing poor performing managers increase within-fund network centrality. The former has a stronger and lasting positive effect up to three quarters ahead. Hiring new fund managers improves bond funds' information access. Fear of forced liquidation, funds perform worse compared to their peers tend to spend more efforts on information collection. It thus results in higher future network centrality. Moreover, higher fund flow predicts higher within fund network centrality. Higher fund flows enlarge funds' capacity to act upon trading information as they can reallocate their portfolio quickly with new cash inflow. It motivates them to access to more information and thus increases their future network centrality.

Considering a time-series analysis, it is found that an existing informational advantage becomes augmented during times of heightened market uncertainty. When there is widespread distress in the financial market, accurate information for trade decisions become more critical than usual. Based on this reasoning, I observe a stronger positive relationship between network centrality and fund performance in periods of higher market uncertainty. I use both crisis periods and the Financial Stress Index from Federal Reserve Bank of St. Louis which combines 18 series including VIX and interest rate spreads, as a proxy for periods of uncertainty. The stronger positive sensitivity remains robust and significant in both specifications.

From a cross-sectional perspective, funds with higher liquid assets can adjust their portfolio at a lower transaction cost. I show a stronger relationship between network centrality and fund performance for funds with more liquid assets. Controlling for network centrality, bond funds with more liquid assets profit more according to the information they receive. They can either sell the liquid assets or use cash holdings to conduct asset reallocation, thereby avoiding high transaction cost induced by selling illiquid bonds.

In order to show that network central bond funds indeed capture profitable information which explains their high future fund performance, I further construct information network by tracking the same trades following through quarters. This enables me to examine herding behavior in bonds funds and differentiate funds who act as a first mover for given information between quarters. Using lead-lag trades between quarters, I build the herding information network with the directional edge pointing from follower to leader. Funds highly centered in this network herd more with most trades similar to other funds in previous quarters. I find that leader funds which lead other bond mutual funds by one quarter ahead predicts 1.2% increases in their future fund alpha. Leader funds which lead other funds by two quarter ahead in trading predict 1.23%. It confirms that the profitable information captured by network central bond funds leads their future high fund performance. To the best of my knowledge, I am the first to construct a trading network using linkage across multiple time periods to empirically measure the pairwise herding behavior from trade level.

The paper is organized as follows. Section 1.2 outlines related literature and contribution. Section 1.3 shows main hypotheses development. Section 1.4 describes weighted information network construction. Section 1.5 presents sample data and variables construction. Section 1.6 discusses the empirical results of the impact of network information on bond fund future returns. Section 1.7 concludes.

1.2 Contribution to Literature

My paper adds to three bodies of knowledge within the finance literature. First, there is growing interest in the inter-connectedness of financial intermediaries and the impact that such connections have on broader financial markets. Current literature focuses primarily on connections within interbank markets from a theoretical perspective (Acemoglu, Vasco, Asuman and Alireza (2012), Di Maggio and Tahbaz-Salehi (2015) and Greenwood, Hanson, and Stein (2015))). There has also been studies performed on the interconnections of non-bank financial intermediaries (Di Maggio, Franzoni, Kermani and Sommavilla (2017a), Di Maggio, Kermani and Song (2017b) on brokers and order flow leakage, Rossi, Blake, Timmermann, Tonks and Wermers (2015) on equity portfolios of UK pension funds linked through same fund managers, Booth, Gurun, and Zhang (2014) on the connection of globally linked institutions and c on flow spillover in same debt asset class). My work makes the unique contribution of identifying information inter-connectedness between corporate bond mutual funds inferred directly from their trading behavior. My constructed information network helps us to investigate possible impacts on pricing efficiency and liquidity in the corporate bond market.

Secondly, I contribute to the empirical literature on herding behavior in the corporate bond market. Cai, Song, Dan and Yi (2017) are the first to study herding in bond markets and focused on how herding destabilizes bond prices. They identify herding intensity by comparing bond's selling/buying volume with the average of all bonds, as commonly done in the equity literature (Brown, Wei, and Wermers (2013)). My network analysis builds upon such work by directly examining each leader and follower pair on every trade and construct time series fund level herding intensity. I provide evidence that herding decreases bond fund risk-adjusted return.

Lastly, my work adds to previous studies on the effect of the network on investment performance. Pareek (2012) links equity mutual funds by their top common holdings to study stock returns and price momentum. Using trades of individual investors from the Turkish stock market, Ozsoylev, Walden, Yavuz and Bildik(2014) find that network centrality predicts high profitability. Rossi, Blake, Timmermann, Tonks and Wermers (2015) present more connected pension fund managers via consultants have higher performance. Ahern (2016) shows stocks linked to central industries perform better. Zhu (2016) studies the connected equity fund managers through their prior career experience. In contrast with their works on the equity market, I provide new evidence on the corporate bond market and the network is constructed directly from the trades from institutional investors. Besides building information networks using information links from the same time period, I also construct herding networks using information linkage across time. Furthermore, my network is weighted to reduce noise in the network construction. In this way, I shed light on herding behaviors and the underlying corporate bond liquidity.

1.3 Main Hypothesis Development

Corporate bonds currently trade over the counter, leading to lower information efficiency relative to the equities, due to the opacity of trades. The higher search cost and information lost (Harris (2015) and O'Hara, Wang, and Zhou (2016)) can be primarily attributed to the lack of a centralized trading platform. Such high search cost augments the value of information in the corporate bond market. The institutional backdrop of the corporate bond market and corresponding mutual fund industry lead us to the following four hypotheses in this section.

Before going into my hypotheses, I clarify a few definitions. Funds always refer to corporate bond mutual funds in this paper. Fund network centrality refers to how central a mutual fund is according to the measure of eigenvector centrality. A bond fund with high network centrality is said to have an inherent information advantage relative to other funds with lower fund centrality (as described in later sections). Fund performance refers to the risk-adjusted return constructed following Goldstein, Jiang, and Ng (2015) and Cutura, Parise, and Schrimpf (2017). The information network described refers to the graph construction in this paper with each note representing a bond fund. The link between nodes depends on if two funds access to the same information evidenced by the same trades within the same time period. It will be described in more detail in subsequent sections.

Hypothesis I: High fund network centrality leads to high bond fund returns

Investors in the central position of the network have increasing informational advantage as they are more densely connected and have more linkages to other investors in the network. They also have earlier access to information signal (Jackson (2009), Ozsoylev and Walden (2011) and Walden (2014)). The high informational advantage thus enables investors to either make more informed investment decisions or act as a first-mover on the information. Both cases lead to higher profitability. Similarly, it can be envisioned that centrally positioned corporate bond mutual funds have greater information advantage which enables them to have higher future fund performance compared to those less connected corporate bond mutual funds in the periphery of the information network. Due to lower trading liquidity relative to equities, the positive relationship between network centrality and bond fund return is likely to be also relatively more persistent and significant.

Hypothesis II: Fund turnover, ranking pressure, and past fund flow influence within-fund time-varying network centrality.

Hiring new managers and firing poor performing managers might improve bond mutual funds' access to information and thus increase within-fund network centrality. Due to competitive pressure, funds performing worse than their peers tend to spend more efforts on information collection. It leads to higher future network centrality. Moreover, since higher fund flows improve bond funds' capacity to trade on profitable information, it motivates them to improve their access to information. Therefore, it increases their future network centrality.

Hypothesis III: The magnitude of Hypothesis I is augmented during market crises.

When market uncertainty level heightens, accurate information is valued more substantially relative to times of low volatility. It thus follows that the positive relationship described in hypothesis I is stronger during periods of market uncertainty.

Hypothesis IV: The magnitude of Hypothesis I is stronger for funds with greater liquidity.

The correlation between network centrality and future fund return is stronger for funds with more liquid assets, since they can execute with a lower cost irrespective of their information advantages. Despite having a high informational advantage, the level of future performance achieved also depends heavily on how efficiently they can reallocate assets in line with their information. Corporate bond mutual funds with more liquid assets are able to adjust current holdings at a lower cost. They can either sell the liquid assets or use cash holding to conduct asset allocation while avoiding high transaction costs. Therefore, the positive relationship between network centrality and fund performance is stronger for funds with more liquid assets.

1.4 Weighted Information Network Construction

1.4.1 Weighted Information Network

This section I present the procedure to construct a weighted information network for corporate bond mutual funds based on their trading behaviors. I obtain quarterly bond level transaction data for each bond funds from eMaxx as described later in the data section. For every quarter, I construct a trading network. Each node in the network represents a bond fund. An edge between any two nodes is formed if two bond funds trade on the same corporate bond with the same direction, either buy or sell, within the same quarter. Bond funds with the same trades tend are assumed to have access to the same information or share the view and hence it forms information linkage in the network. This is built on Ozsoylev, Walden, Yavuz and Bildik (2014) which construct the network of individual investors in Turkish stock exchange using the same stock trades. Improving upon their unweighted information network, I construct a weighted information network of corporate bond funds.

I assign the weight to each edge linking two funds in the network based on pairwise quarterly overlapping trading intensity. It is measured by aggregating all the same bonds both funds trade each quarter scaled by their total quarterly trading quantities. The measurement is constructed as:

$$EdgeWeight_{i,j,t} = \frac{\frac{\sum_k Trade_{i,t,k}}{Q_{i,t}} + \frac{\sum_k Trade_{j,t,k}}{Q_{j,t}}}{2}$$

where $EdgeWeight_{i,j,t}$is the weight of the edge between bond fund *i* and *j* in quarter *t*. $Q_{i,t}$ represents total trading volume by bond fund *i* in quarter *t*. $Trade_{i,t,k}$ is the trading quantity on bond k by bond fund i and k stands for the same bond traded by both mutual funds during the quarter t. The higher percentage of one fund's quarterly trading overlaps with the other bond fund, the stronger is the information linkage between these two bond funds. Hence, in my bond mutual fund trading network, the magnitude of edge weight captures the intensity of correlated trading and the strength of pairwise information linkage.

The weighted trading network I design above has two advantages. First, by capturing the intensity of pairwise bond funds' trade overlap as edge weights, my network is more informative than the unweighted trading network by Ozsoylev, Walden, Yavuz and Bildik (2014) and Sung (2016). Edges in their networks only capture that two investors have the same stock stocks but fail to differentiate edge linkages with different trading overlap intensities. This could emphasize edges of thin trades disproportionally and include more noise in the network. Second, my network does not capture the stale information caused by frictions to sell illiquid corporate bonds. Pareek (2012) adopts common equity holdings to form linkages which might record the information change with a lag. If two bond funds have different views on the illiquid bond they commonly hold. One of the funds plans to sell the bond but is unable to liquid the bond positions due to high transactions cost this quarter, and sells the bond next quarter instead. Network based on common holdings still connects two bond funds via this hard-to-sell illiquid bond and doesn't reflect the change of information until next quarter. Contrarily, my network using trading behavior to form the linkage overcomes this concern and records no connection between tow bond funds. It decreases the change to record the stale information in the network. My network serves well as an effective tool to analyze information linkage among mutual funds via their correlated trading behavior.

Besides the weighted information network, I construct another directed trading networks. The directed buy or sell trading information network is built using only the buy or sell trades. Separating the buy trades and the sell trades allows me to take into consideration the direction in the network and measure the differential effect brought by buy/sell information on bond markets. The difference between buy network and sell network enriches my understanding of sources of information and their differences in market dynamics.

1.4.2 Eigenvector Centrality Measurement

To measure how central a bond fund in the network, I adopt eigenvector centrality measurement (Bonacich (2007)). This measure considers both the number of connections each node links to and the importance and connectivity of nodes it connects to. The more densely connected nodes one node connects to weights more than the scarcely connected nodes one node links to. The eigenvector centrality measurement is thus recommended in literature as a more superior measure to capture the information diffusion (Ozsoylev, Walden, Yavuz and Bildik (2014) and Walden (2014)).

Eigenvector centrality of each node is proportional to the aggregate sum of the eigenvector centralities of nodes connected to it. It is formulated as beloved.

$$EiC_{i,t} = \frac{1}{\delta}\sum_{j:j\neq i} L_{i,j} * EiC_{j,t}$$

$$\delta * EiC = L * EiC$$

where $EiC_{i,t}$ is eigenvector centrality of bond fund i in the network at quarter t. $L_{i,j}$ equals to 1 if bond fund i links to bond fund j or 0 otherwise. δ is a constant which denotes the eigenvalue of matrix L with EiC as its eigenvector. The power iteration applies in order to solve systems of equations. The eigenvector associated with the biggest eigenvalue becomes network centrality eigenvector (Bonacich (2007)). Eigenvector centrality assigns importance of each connection by its eigenvalue. Walden (2014) theoretically characterizes the close relationship between eigenvector centrality with information advantages of investors positioned in the network. The central bond funds receive more informative signals and also receive information at shorter delays compared to periphery bond funds in the network.

1.5 Data Construction

My primary data on corporate bond mutual funds come from five databases: Thomson Reuters Lipper eMaxx, Fixed Investment Securities Database (FISD), Enhanced Historic Trade Reporting and Compliance Engine (TRACE Enhance), CRSP Survivorship Bias Free Mutual Fund Database and Datastream. My sample period covers from 2004 Q2 to 2014 Q4 on a quarterly frequency.

1.5.1 Corporate bond quarterly holding and trading data

Thomson Reuters eMaxx database provides comprehensive end of quarter security level information on corporate bond holdings and trading positions by various institutional investors including mutual funds, pension funds, and insurance companies. In this paper, I focus on holdings and trading by corporate bond mutual domiciled in the U.S. Importantly, the eMaxx database contains all defunct funds and free of both survivorship bias and reporting biases. Literature (for example, Massa and Zhang (2012), Becker and Ivashina (2015), Choi and Kronlund (2017) and Cai, Song, Dan and Yi (2017)) has used eMaxx for corporate bond market studies.

1.5.2 Corporate bond pricing related data

I merge eMaxx with FISD by security level eight-digit CUSIP to obtain detailed information on bond issuer level and bond issue level characteristics (time to maturity, etc.). After retrieving complete historical ratings for each corporate bond, I use the highest rating of three major credit rating agencies (S&P, Moody's and Fitch) as the final bond rating each quarter. In order to access to pricing related information, I further merge the eMaxx with TRACE Enhance database which contains complete over the counter transaction data from 2002 to 2014. Equipped with the high frequency transaction data, I am able to construct various liquidity and bond return measurements. I use CRSP bond return data as an additional source. I adopt Dick-Nielsen (2014) methodology to clean TRACE Enhance database which corrects for liquidity biases, errors and double counted transactions problems in TRACE data.

1.5.3 Corporate bond mutual fund level characteristics

I augment the eMaxx database with CRSP mutual fund database to gather fund level characteristics. I first use CRSP Lipper Object code to filter all mutual funds in eMaxx to obtain bond funds primarily investing in corporate bonds. I further exclude index fund and exchange-traded funds. CRSP and eMaxx don't share the same fund identification code. Therefore, I match funds in these two databases by fund legal name to construct the fund level link. I utilize the fuzzy matching algorithm and searching algorithm from computer science literature to conduct the first-round matching and then I manually check each name to guarantee the accuracy. This results in 3006 unique fund share classes from CRSP with 839 funds in eMaxx.

Equipped with this fund level link, I retrieve various characteristics of mutual funds from CRSP including mutual fund investment style, total assets, returns, flow, investors and cash holdings. Since fund share classes from the same fund hold the same portfolio and differ by expense ratios and fee structures, I aggregate all variables from CRSP share classes to fund level using value weighted method with the total net asset from the previous period as weight.

1.5.4 Market return and bond mutual fund benchmark index return

Lastly, I download various mutual fund benchmarks including Vanguard Total Bond Market Index Fund return from DataStream (Ng, Wang, and Wang(2015)) and aggregate market excess return from CRSP Fama French Factors Library. My final sample includes 22301 bond mutual fund-quarter observations for fund level analysis and 92,251 bond issue-quarter observations for issue level analysis.

1.5.5 Variables construction

Consistent with related literature on corporate bond mutual fund, I compute corporate bond mutual fund risk adjusted return, alpha, following Goldstein, Jiang, and Ng (2017) and Cutura, Parise, and Schrimpf (2017). I perform 12-month window rolling regressions of monthly bond fund excess return on aggregate stock market excess return obtained from CRSP and aggregate total bond market index excess return. Alpha is the intercept of the regression. Barclays U.S. Aggregate Index Return is used as a proxy for aggregate total bond market returns. This procedure becomes a common practice to compute bond fund alpha due to two reasons. First, Barber, Huang, and Odean (2016) and Berk and van Binsbergen (2016) shows that investor does not use the sophisticated model to adjust for risk except the market benchmark return. Second, choice of different model to adjust for bond fund risk does not change the relative ranking of bond funds (Cutura, Parise, and Schrimpf (2017) and Blake, Elton, and Gruber (1993)). In addition, I include default premium risk factor and the term premium risk factor to measure the risk adjusted return (Ng, Wang, and Wang (2015)). Default premium is measured as the difference of returns between high yield bond index and the intermediate government bond. The term premium is calculated as the return difference between intermediate government bond index and short-term government bond index.

Quarterly corporate bond fund flow is computed as:

$$FundFlow_{i,t} = \frac{TotalNetAssets_{i,t} - TotalNetAsset_{i,t-1} * (1 + R_{i,t})}{TotalNetAsset_{i,t-1}}$$

where $R_{i,t}$is the net accumulative return of fund *i* during quarter t. $FundFlow_{i,t}$ is fund flow of bond fund *i* at quarter t. This method is the standard in mutual fund flow literature (Chen, Goldstein, and Jiang (2010)).

I compute corporate bond level liquidity using various liquidity measurement. The first one is Amihud (2002)'s price impact measure. It measures the ratio of price movement over the dollar trading volume. It is defined as

$$BondLiquidity_{i,D} = \frac{1}{N_{i,D}} \sum_{k} \frac{\frac{|Price_{i,D}^{k} - \ Price_{i,D}^{k-1}|}{Price_{i,D}^{k-1}}}{TradeQuantity_{i,D}^{k}}$$

where $Price_{i,D}^{k}$ represents the price of *k*-th trade on bond *i* at day *D* during quarter t. $N_{i,D}$ is the total number of trades on bond *i* on day D.

The second liquidity measurement is Lesmond, Ogden, and Trzcinka (1999)'s zero trading days. It measures the stagnancy of bonds which do not trade for a long period. The longer the zero trading days a bond has, the less liquid the bond is.

$$ZeroTradingDays_{i,t} = \frac{\text{number of zero return days}_{i,t} + number\ of\ zero\ trade\ days_{i,t}}{total\ number\ of\ trade\ days_t} * 100$$

In addition, the corporate bond level liquidity is measured using Lesmond, Ogden, and Trzcinka (1999)'s imputed roundtrip cost. It measures the transaction cost using bond trades executed within the short window and with the same trading volume.

$$ImputedRC_{i,t} = \frac{Price_{max,i,t} - Price_{min,i,t}}{Price_{min,i,t}} * 100$$

Another bond liquidity measure is Dick-Nielsen, Feldhütter, and Lando (2012)'s the interquartile price range. It is calculated as the difference between the interquartile prices of trade bonds. Lastly, the average trade spread is adopted to measure bond level liquidity which is the quarterly average of bid-ask spreads for each bond.

Table 1.1 presents the summary statistics of the main variables I include in my analysis.

[Insert Table 1.1 here]

1.6 Empirical Analysis: Impact of Information Networks on Bond Mutual Funds

In this section, I present that corporate bond funds posited in the central of the trading network have higher future fund performance. I start with the panel fixed effect model and then use a portfolio analysis to examine the predictive power of network centrality on bond future performance. Moreover, I evaluate the within-fund time varying changes of eigenvector centrality resulting from fund manager turnover, competitive/ranking pressure, and fund flow fluctuations. I conclude with the potential mechanism via which the positive relationship between network centrality and bond return is formed.

1.6.1 Positive relationship between eigenvector network centrality and bond fund returns

Using fixed effect panel regression model, I examine the effect of network centrality on future bond fund returns.

$$FundReturn_{i,t+1} = \beta_1 NetworkCentrality_{i,t} + Controls_{i,t} + FixedEffect_i + \epsilon_{i,t}$$

where $FundReturn_{i,t+1}$ is corporate bond fund future return. I use four sets of bond fund returns: one month ahead net return, 3 month ahead net returns, one year ahead net returns and risk adjusted fund return $FundAlpha_{i,t+1}$. $NetworkCentrality_{i,t}$ is eigenvector centrality for bond fund *i* at quarter t. High eigenvector centrality indicates a bond fund in a more central position of the network with high informational advantages. I include the logarithm format of fund total net assets, quarterly trading volume, number of bonds traded and fund fixed effects to control for fund characteristics that might correlate with network centrality. I want to measure the influence of over time changes in network centralities on bond fund future returns.

[Insert Table 1.2 here]

I am interested in β_1 which measures the sensitivity between eigenvector centrality and bond fund performance. Table 1.2 reports the estimation results. Consistent with the theoretical prediction, I find that bond mutual funds with high network centrality lead to a high future return in all specifications. In particular, the positive relationship remains for up to one ahead return and risked adjusted fund alpha. More specifically, one standard deviation increase in the eigenvector centrality leads to 0.16% increase the corporate bond fund's one month ahead return, 0.39% in three-month ahead return, 1.29% in one year return, and 0.33% increase in annualized alpha. This indicates that a positive relationship between network centrality and fund return is persistent and stronger for the long-term fund returns. It suggests that high network-central bond funds indeed have an increasing informational advantage which enables them to make informed trading decisions and thus results in high future fund returns. This lends support to my first hypothesis.

[Insert Table 1.3 here]

In addition, I use the portfolio-based analysis to examine the impact of network centrality on bond fund future performance. I sort bond mutual funds into five portfolios according to their eigenvector centrality at the end of each quarter. I hold these portfolios for six months and then rebalance them. Table 1.3 presents the annualized alpha of the portfolio results. Bond mutual funds with high eigenvector centrality in the High portfolio outperform the bond funds with low eigenvector centrality by 1.35% per year in their risk adjusted future fund return. Funds in the lower decile experience decrease in their future risk adjusted performance. This finding speaks in favor of network centrality leading to high future fund returns due to informational advantages from being in the center of the information network.

[Insert Table 1.4 here]

I further rule out the alternative story of style investing. Corporate bond funds in the same investment style with the same benchmark index tend to trade on similar corporate bonds. Some might concern that the positive relationship between network centrality and corporate bond fund performance results from style investing rather than being central and informative in the information network. I rule out this possibility by constructing the network using only trades sharing by different styles of corporate bond mutual fund to control for this style effects. I find that positive relationship between network centrality and fund returns still remains significant for within styles and between different styles as shown in Table 1.4. In addition, if I include style as a control instead of constructing subnetwork described as above. The positive relationship remains. Therefore, both pieces of evidence support my hypothesis also try the style adjusted return, the result remains the same. Therefore, both pieces of evidence support my hypothesis that informational advantage resulting from being center in the network leads to higher fund future returns.

[Insert Table 1.5 here]

Moreover, I construct an information network by using buy trades and sell trades separately. It allows me to analyze if the information on buying and information on selling would have different effects on bond future returns. I find that information on selling is 50% more valuable than information on buying in terms of future risk adjusted fund alpha. However, information on what to buy is more profitable with respect to short term one month net return, as shown in Table 1.5.

1.6.2 Within-fund time varying changes of eigenvector centrality

This section I assess how within-fund eigenvector centrality changes over time due to manager turnover, competitive pressure, and fund flow fluctuation. This analysis is crucial as the incentive to improve fund network centrality directly speaking about bond fund managers' effort to acquire the information.

1.6.2.1 Time varying changes in eigenvector centrality due to manager turnover

Hiring new managers or firing poor performing managers increases within-fund future eigenvector centrality. Specifically, I run the following fixed effects regression:

$$NetworkCentrality_{i,t+1\rightarrow t+3} = \beta_1 FireManager_{i,t} or\ HireNewManager_{i,t} + Controls_{i,t} + FixedEffect_i + \epsilon_{i,t}$$

where $FireManager_{i,t}$ denotes events when a fund manager is fired during the quarter *t*. $HireNewManager_{i,t}$ represents the events when a new fund manager is hired and there is no existing fund managers is fired during the same quarter *t*. Both variable measure the fund manager turnover. I examine the eigenvector centrality in various look ahead period from quarter *t+1* to quarter *t+3*.

[Insert Table 1.6 here]

I am interested in $\beta_1$which measures the sensitivity between fund manager turnover and within fund future eigenvector centrality. The estimated coefficients reported in Table 1.6 shows that both hiring new manager and firing poor managers increase within-fund future eigenvector centrality. However, fund turnover from hiring new manager has stronger and more lasting effects the turnover coming from firing existing managers.

The coefficients of $FireManager_{i,t}$ are positive during look ahead quarters but the positive effect is fading gradually. It is no longer statistically significant after two quarters. In contrast, the positive effect on network centrality resulting from hiring new managers gets strong over time. The impact of hiring new managers is more than two times larger than that of firing existing poor performing ones.

1.6.2.2 Time varying changes in eigenvector centrality due to ranking pressure

Due to competitive pressure, funds perform poorly in the past has the incentive to spend more effects to gather information to improve their fund performance. It thus might increase their future eigenvector centralities, which is a proxy for their informational advantage.

$$NetworkCentrality_{i,t+1} = \beta_1 LowRankedFund_{i,t} + \beta_2 HighRankedFund_{i,t} + Controls_{i,t} + FixedEffect_i + \epsilon_{i,t}$$

The above model is conducted. Bond mutual funds are classified into quantiles based on their past 12-month fund returns. $LowRankedFund_{i,t}$ denotes funds in the bottom quantile which perform the worst compared to their peers. $Hi\mathrm{ghRankedFund}_{\mathrm{i,t}}$ denotes funds in the top quantile which rank the highest in terms of their fund performance.

[Insert Table 1.7 here]

Positive and statistically significant coefficient β_1 in Table 1.7 shows that funds ranked lower among their peers are motivated to improve their future network centrality. In contrast, $\beta_2$is not statistically significant. It shows that bond funds perform wells with high rankings are less likely to increase their within-fund network centrality. This is consistent with the predictions in Cutura, Parise, and Schrimpf (2017) and Goldstein, Jiang, and Ng (2017). Both show that the worst performing bond funds are afraid of losing the tournament game which would lead to fund runs and forced liquidation. As results, those worst performing bonds are willing to make more efforts and take more risk hoping to gamble for resurrection. As they make more efforts on information gathering, it increases their future network centrality.

1.6.2.3 Time varying changes in eigenvector centrality due to fund flow fluctuations

Within funds, higher fund flows predict increases in future eigenvector centrality as funds have a higher capacity to reallocate their portfolios in response to new trading information. As shown in Table 1.8, I find that one standard deviation increase in past fund flow increases within fund eigenvector centrality.

[Insert Table 1.8 here]

1.6.3 Potential mechanism behind the positive relationship between centrality and fund return

I hypotheses that positive relationship between eigenvector centrality and fund performance is formed via the mechanism that high network central funds indeed capture profitable information. The profitable information enables bond funds to make a wise decision which leads to a boost in the future bond fund returns. In order to directly capture that the information that network central is profitable, I construct a leader-follower herding network. The herding network allows me to track the same trading decision over time and measure the profitability of funds who access to this information first.

The herding network is constructed using the same trading behavior following across quarters, rather than using the same trades executed within the same quarters in the networks I show previously. For each trade placed by the bond fund during a given quarter t, I keep track of the same trades occurring during quarter t+1 and quarter t+2. This herding information network allows us to study the herding behavior across time periods at a leader-follower pairwise trade level, rather than at an aggregate level as done in the current literature. Based on the number of leading legs, bond funds are then classified into quantiles. Leaders funds are funds in the top quantile with top numbers of leading leads which means they are followed by many other bond funds across quarters. Similarly, Follower funds are funds in the bottom quantile which herd the most. Specifically, I run the following specification.

$$NetworkCentrality_{i,t+1} = \beta_1 Leader_{i,t} + \beta_2 Follower_{i,t} + Controls_{i,t} + FixedEffect_i + \epsilon_{i,t}$$

[Insert Table 1.9 here]

If the coefficient of leader fund is positive, that means that the information collected by leaders in making their trades is profitable information. From the coefficient report in Table 1.9, I find that leader funds which lead a quarter ahead predicts 1.2% increase their alpha. Moreover, for leader funds which lead other funds by two quarter ahead in trading predicts 1.23% a slightly higher increase in their risk adjusted return per year. This evidence supports that the information captured in the trading network is profitable information which serves as a potential channel via which the positive relationship between network centrality and bond fund future returns is formed. By being in the central positions of a trading network, network central funds tend to capture more profitable information. Hence, profitable information enables to network central funds to have higher fund returns in the future.

1.6.4 Positive relationship between network centrality and bond fund performance for corporate bond funds is stronger during the period of crisis and uncertainty.

$$FundAlpha_{i,t+1} = \beta_1 NetworkCentrality_{i,t} + \beta_2 NetworkCentrality_{i,t} * Crisis_t$$
$$+\beta_3 Crisis_t + Controls_{i,t} + FixedEffect_i + \epsilon_{i,t}$$

where $Crisis_t$ indicates periods of market uncertainty. I use two measures to proxy for uncertainty periods. One is the crisis dummy variable including the past financial crisis. The other one is Financial Stress Index released by Federal Bank of St. Louis which measures the degree of financial stress and market illiquidity constructed from 18 weekly series including various interest rates changes, VIX and TED spreads.

Table 1.10 presents my findings. From Table 1.10 column 1 and column 2, I find that eigenvector centrality predicts stronger risk adjusted returns during the periods of financial crisis or financial market stress. This result is in line with my hypothesis two that information is more valued and critical during high market uncertain and stressed time. Hence, it helps high network central bond funds profit more from their information advantage.

[Insert Table 1.10 here]

1.6.5 The sensitivity between network centrality and future fund return is stronger for corporate bond mutual fund with high fund level liquidity.

$$
\begin{aligned}
FundAlpha_{i,t+1} &= \beta_1 NetworkCentrality_{i,t} + \beta_2 NetworkCentrality_{i,t} \\
&\quad * FundLiquidity_{i,t} \\
&\quad + \beta_3 FundLiquidity_{i,t} + Controls_{i,t} + FixedEffect_t + \epsilon_{i,t}
\end{aligned}
$$

where $FundLiquidity_{i,t}$ denotes bond mutual fund level liquidity. I adopt two proxies for fund level liquidity: the percentage of cash holding and the aggregate liquidity of corporate bonds holdings by bond funds each quarter.

Corporate bond funds with more liquid assets are able to act on new information with a lower asset reallocation cost. They can avoid selling illiquid holdings at unfavorable times or high transaction cost and use cash or liquid assets to accomplish portfolio adjustment. Therefore, controlling for network centrality, bond funds with high liquid assets predicts a higher future fund return. Results in Table 1.10 column3 and column 4 lends support to this claim and shows that bond with more liquid assets enjoys an additional 1% to 4% increase in future fund returns.

1.7 Conclusion

This paper investigates the influence of the information network of corporate bond mutual funds on fund performance, herding behaviors, and corporate bond markets. The Information network of bond mutual funds is formed by bond funds executing same bond trades with the same direction within the same quarter. The network-central bond funds have the informational advantage which leads to higher fund future returns, both net returns and risk adjusted returns, even after controlling for style investing. The positive relationship between network centrality and fund performance is stronger during periods of stressed time, as information is more valued during those turbulent times, and for funds with more liquid assets as they can act upon the information with a lower asset allocation cost. Moreover, I exploit within-fund time varying changes in network centrality due to manager turnover, ranking pressure, and fund flow fluctuation. By constructing the herding information network, I further show that network central bond funds capture profitable information. This paper makes an important contribution to the literature by identifying information flow and inter-connectedness between corporate bond mutual funds directly from their actual trades.

1.8 Figures and Tables

Table 1.1: Summary Statistics

This table presents the fund level characteristics of active corporate bond mutual funds studied in my sample from 2004 Q2 to 2014 Q4 and the bond level characteristics of corporate bonds traded by these corporate bond mutual funds during the same period. All observations are available on a quarterly basis. I report the number of observations, mean, standard deviation, minimum value and maximum value of each variable. Eigenvector Centrality measures the centrality of each fund in the information network based on both the number of connections the fund has within the quarter and how important and connected other funds this fund links to (Bonacich (2007)). Alpha is the risk adjusted fund performance which is calculated by running 12-month rolling regression of bond fund excess net return on the aggregate stock market excess return from CRSP and the aggregate Vanguard total bond market index fund excess return, following Goldstein, Jiang, and Ng(2017). Alpha reported in the paper is the annualized version of monthly alpha. 1-month, 3-month and 12-month are bond fund net return for one month up to one year. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Quarterly Fund Flow is the percentage of fund low in a given quarter. Fund Holdings liquidity is the fund level liquidity. It is calculated by aggregating the Amihud illiquidity measure (Amihud (2002)) of each compote bond in fund holdings with holdings as weight. I multiply this aggregated number by -1 to show fund level liquidity in the regression. Cash Equivalents Holdings is the percentage of cash and government bonds held by funds in a given quarter. Eigenvector Centrality Weighted Bond Ownership is constructed at the bond level in each given quarter. For every corporate bond ever traded by any corporate bond funds, I multiply the percentage of the bond held by each corporate bond fund with its corresponding bond fund network centrality and sum this number across all the bond funds who hold this bond in the given quarter. This variable measures the intensity of a corporate bond owned by high network centrality bond fund. Bond Rating is the max of ratings coming from three credit rating agencies: Moody's, S&P and Fitch. Maturity is time to maturity. Average Trade Spread is the quarterly average bid-ask price for each bond. Quarterly Bond Return is quarterly return for each bond. Amihud Illiquidity is the bond level illiquidity measurement (Amihud (2002)). Log Trading Volume is the logarithm of the bond level total quarterly trading volume. Financial Stress Index is obtained from St. Lois Fed which captures the degree of financial stress and market uncertainty in the financial markets. It is calculated from 18 macroeconomics series including VIX, seven interest rate series, six yield spread series and other indicators. It is a widely used proxy for market situations and crisis periods. Stars represent statistical significance. (***p<0.001, **p<0.005, *p<0.1, respectively)

Fund Level	N	Mean	St. Dev.
Eigenvector Network Centrality	21,330	0.18	0.20
12-month Alpha	21,502	0.03	0.14
1-month Return	21,415	0.01	0.02
3-month Return	21,445	0.01	0.03
12-month Return	21,502	0.04	0.08
Log Total Net Assets	21,397	5.71	1.82
Log Trading Volume	21,330	3.48	1.37
Log Traded Bond Issues	21,330	13.16	3.20
Quarterly Fund Flow	20,902	0.02	0.12
Fund Holdings liquidity	19,200	0.07	0.06
Cash Equivalents Holdings	20,524	10.55	14.30

Bond Level	N	Mean	St. Dev.
Eigenvector Centrality Weighted Bond Ownership	90,310	0.88	0.93
Bond Rating	92,251	8.62	3.17
Maturity	92,251	111.06	108.27
Amihud Illiquidity	90,405	0.19	0.23
Average Trade Spread	86,939	0.62	0.48
Quarterly Bond Return	89,739	0.02	0.06
Log Trading Volume	88,180	18.02	1.35

Macroeconomics Series	N	Mean	St. Dev.
Financial Stress Index	22,341	-0.27	1.13

Table 1.2: Relationship between Network Centrality and Bond Mutual Fund Future Returns

This table presents the relationship between network centrality and corporate bond fund future performance in the aggregate information network. The underlying information network is constructed with nodes representing each bond mutual funds and linkage between any nodes is formed if any two bond funds have the same trades (either sell or buy) during the same quarter. The weight of the edge is equal to the percentage of overlapping trades they share out of their corresponding quarterly trading quantities. The weight measures the intensity of overlapping trading behavior between any two mutual funds in a govern quarter. Eigenvector Network Centrality measures the informational advantage a bond fund has based on the number of connections it has and the importance of bond funds it links to. Eigenvector Network Centrality is standardized in regression. Annualized Alpha is future risk adjusted fund performance, which is the annualized version of monthly alpha. Monthly alpha is calculated by running 12-month rolling regression of bond fund excess net return on the aggregate stock market excess return from CRSP and the aggregate Vanguard total bond market index fund excess return, following Goldstein, Jiang, and Ng (2017). I also include additional bond risk factors from Ng, Wang, and Wang (2015). 1-month, 3-month and 12-month are bond fund future net return up to one year. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Fund fixed effects are included in the regression. Stars represent statistical significance. (***p<0.001, **p<0.005, *p<0.1, respectively)

	Aggregate Information Network			
	1-month Return	3-Month Return	12-month Return	Annualized Alpha
	(1)	(2)	(3)	(4)
Eigenvector Network Centrality	0.157***	0.389***	1.293***	0.331**
	(0.024)	(0.040)	(0.093)	(0.146)
Total Net Assets	0.013	-0.300***	-1.598***	-0.455***
	(0.025)	(0.042)	(0.098)	(0.154)
Trading Volume	0.035	0.182***	0.693***	-0.047
	(0.025)	(0.042)	(0.097)	(0.153)
Traded Bond Issue Counts	0.067	-0.232***	-1.588***	0.098
	(0.051)	(0.086)	(0.199)	(0.314)
Fund Fixed Effect	Yes	Yes	Yes	Yes
Observations	20,762	20,777	20,777	20,777
R^2	0.006	0.008	0.025	0.007

Table 1.3: Network Centrality-Fund Performance Relationship: Bond Funds in Low and High Quintiles of Network Centralities

This table presents the relationship between network centrality and corporate bond fund future performance for bond funds with low and high quintiles of network centralities. I classify bond mutual funds into quantiles based on their eigenvector network centrality. Eigenvector Network Centrality measures the informational advantage a bond fund has based on the number of connections it has and the importance of bond funds it links to. Eigenvector Network Centrality is standardized in regression. Annualized alpha is the future risk adjusted fund performance, which is the annualized version of monthly alpha. Monthly alpha is calculated by running 12-month rolling regression of bond fund excess net return on the aggregate stock market excess return from CRSP and the aggregate Vanguard total bond market index fund excess return, following Goldstein, Jiang, and Ng (2017). Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Portfolio Sorted by Eigenvector Network Centrality
	Annualized Alpha
Low	-0.098
	(0.202)
2	0.029
	(0.139)
3	0.223**
	(0.111)
4	0.521***
	(0.143)
High	1.248***
	(0.199)
High-Low	1.350***
	(0.271)

Table 1.4: Network Centrality-Fund Performance Relationship: Information Networks for funds of Same Investment Styles vs Information Networks for Funds of Different Investment Styles

Following two tables in this section present the relationship between network centrality and corporate bond fund future performance in information network for funds with the same investment styles and information network for bond funds across different investment styles. The information networks for funds with different styles is constructed using only the same trades sharing by funds with different investment categories during the same quarter. Similarly, the information networks for funds with the same styles are constructed using the same trades sharing by funds within the same investment categories in each quarter. The investment style is determined by the Thomson Reuters Lipper Object Code. Funds within the same investment category use similar benchmark indexes. The positive relationship between fund centrality and fund return remains significantly even for information network constructed using trades from funds with different investment style. By controlling for investing styles, it addresses the concern that the positive relationship between centrality and fund returns is only driven by styling investing or index following rather than information sharing. In the regression, Eigenvector Network Centrality measures the informational advantage a bond fund has based on the number of connections it has and the importance of bond funds it links to. Eigenvector Network Centrality is standardized in regression. Annualized alpha is the future risk adjusted fund performance, which is the annualized version of monthly alpha. Monthly alpha is calculated by running 12-month rolling regression of bond fund excess net return on the aggregate stock market excess return from CRSP and the aggregate Vanguard total bond market index fund excess return, following Goldstein, Jiang, and Ng (2017). 1-month, 3-month and 12-month are bond fund future net return up to one year. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Fund fixed effects are included in the regression. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Information Network within same investment style funds			
	1-month Return	3-Month Return	12-month Return	Annualized Alpha
	(1)	(2)	(3)	(4)
Eigenvector Network Centrality	0.249***	0.330***	0.930***	0.432*
	(0.038)	(0.065)	(0.152)	(0.256)
Total Net Assets	0.003	-0.320***	-1.656***	-1.209***
	(0.025)	(0.042)	(0.098)	(0.165)
Trading Volume	0.039	0.186***	0.707***	0.122
	(0.025)	(0.042)	(0.098)	(0.165)
Traded Bond Issue Counts	0.069	-0.209**	-1.500***	-0.721**
	(0.051)	(0.086)	(0.200)	(0.338)
Fund Fixed Effect	Yes	Yes	Yes	Yes
Observations	20,762	20,777	20,777	20,777
R^2	0.006	0.005	0.017	0.005

	Information Network between different investment style funds			
	1-month Return	3-Month Return	12-month Return	Annualized Alpha
	(1)	(2)	(3)	(4)
Eigenvector Network Centrality	-0.040**	-0.017	0.338***	0.320**
	(0.020)	(0.034)	(0.078)	(0.132)
Total Net Assets	0.010	-0.312***	-1.656***	-1.217***
	(0.025)	(0.042)	(0.098)	(0.165)
Trading Volume	0.034	0.182***	0.713***	0.133
	(0.025)	(0.042)	(0.098)	(0.165)
Traded Bond Issue Counts	0.095*	-0.185**	-1.548***	-0.792**
	(0.051)	(0.087)	(0.201)	(0.339)
Fud Fixed Effect	Yes	Yes	Yes	Yes
Observations	20,762	20,777	20,777	20,777
R^2	0.004	0.003	0.016	0.005

Table 1.5: Network Centrality-Fund Performance Relationship: Information Networks on Buying vs Information Networks on Selling

Following two tables in this section present the relationship between network centrality and corporate bond fund future performance in the information network on buying and information network on selling separately. The underlying information network on buying is constructed similar to the aggregate information network presented in table 2, except that only buying trades are used to construct the information network. Nodes represent each bond mutual funds in the network. However, two bond funds with same buy trades during the same quarter form an edge link in the network. Similarly, for the information network on selling, only bond funds share the same sell trades form a pairwise link between nodes. Different from aggregate information network using all the trades, by separating buy and sell information I am able to further capture the differential information content. Information on selling opportunities results in 50% higher risk adjusted returns. In the regression, Eigenvector Network Centrality measures the informational advantage a bond fund has based on the number of connections it has and the importance of bond funds it links to. Eigenvector Network Centrality is used in a standardized form in regression. Annualized alpha is future risk adjusted fund performance, which is the annualized version of monthly alpha. Monthly alpha is calculated by running 12-month rolling regression of bond fund excess net return on the aggregate stock market excess return from CRSP and the aggregate Vanguard total bond market index fund excess return, following Goldstein, Jiang, and Ng (2017). 1-month, 3-month and 12-month are bond fund future net return up to one year. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of a number of traded bond issues by each fund quarterly. Fund fixed effects are included. Stars represent statistical significance. (***p<0.001, **p<0.005, *p<0.1, respectively)

	Information Network on Buying Information			
	1-month Return	3-Month Return	12-month Return	Annualized Alpha
	(1)	(2)	(3)	(4)
Eigenvector Network Centrality	0.102***	0.163***	0.308***	0.168**
	(0.019)	(0.033)	(0.077)	(0.070)
Total Net Assets	-0.014	-0.347***	-1.681***	-1.275***
	(0.026)	(0.044)	(0.103)	(0.094)
Trading Volume	0.051*	0.209***	0.726***	0.182*
	(0.027)	(0.045)	(0.106)	(0.097)
Traded Bond Issue Counts	0.026	-0.304***	-1.710***	-0.913***
	(0.055)	(0.092)	(0.215)	(0.196)
Fund Fixed Effect	Yes	Yes	Yes	Yes
Observations	20,098	20,113	20,113	20,113
R^2	0.005	0.005	0.018	0.019

	Information Network on Selling Information			
	1-month Return	3-Month Return	12-month Return	Annualized Alpha
	(1)	(2)	(3)	(4)
Eigenvector Network Centrality	-0.013	-0.138***	0.225***	0.409***
	(0.019)	(0.032)	(0.076)	(0.130)
Total Net Assets	-0.009	-0.364***	-1.848***	-1.304***
	(0.028)	(0.047)	(0.109)	(0.186)
Trading Volume	0.047*	0.206***	0.847***	0.188
	(0.027)	(0.046)	(0.108)	(0.183)
Traded Bond Issue Counts	0.078	-0.179*	-1.818***	-1.006***
	(0.056)	(0.095)	(0.223)	(0.378)
Fund Fixed Effect	Yes	Yes	Yes	Yes
Observations	19,733	19,740	19,740	19,740
R^2	0.004	0.005	0.017	0.006

Table 1.6: Within Fund Time Varying Changes of Network Centrality: Manager Turnover

This table presents the within-find time varying change of eigenvector network centrality due to fund manager turnover. Event Fire Manager happens when a fund manager is fired. Event Hire New Manager happens when a new fund manager is hired, and no existing managers are fired during the same quarter. Time t is the timing when the event happens. Time t+1 to time t+3 are the lookahead periods. Eigenvector Network Centrality measures the informational advantage a bond fund has based on the number of connections it has and the importance of bond funds it links to. Eigenvector Network Centrality is standardized in regression. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Fund fixed effects are included. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Network Centrality							
	t	t+1	t+2	t+3	t	t+1	t+2	t+3
Fire Manager	0.028***	0.023**	0.017*	0.003				
	(0.010)	(0.010)	(0.010)	(0.011)				
Hire New Manager					-0.014	0.056**	0.062**	0.065**
					(0.026)	(0.027)	(0.027)	(0.027)
Total Net Assets	-0.041***	-0.023***	-0.021**	-0.023***	-0.039***	-0.023***	-0.021**	-0.022***
	(0.008)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)
Trading Volume	0.011	-0.005	-0.008	-0.007	0.010	-0.005	-0.007	-0.007
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Traded Bond Issue Counts	0.092***	0.052***	0.033*	0.004	0.096***	0.051***	0.032*	0.003
	(0.016)	(0.017)	(0.017)	(0.017)	(0.016)	(0.017)	(0.017)	(0.017)
Fund Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,057	19,022	18,270	17,564	19,340	19,022	18,270	17,564
R^2	0.018	0.003	0.001	0.001	0.019	0.003	0.001	0.002

Table 1.7: Within Fund Time Varying Changes of Network Centrality: Competitive/Ranking Pressure

This table presents the within-find time varying change of eigenvector network centrality due to peer or ranking pressure. Event Fire Manager happens when a fund manager is fired. Each quarter, I classify bond mutual funds into quantiles based on their past 12-month fund performance. High Ranked Funds are funds in the top quantile which outperform their peers in the past 12 months. Low Ranked Funds are the funds in the bottom quantile which ranks lowly among their peers. Eigenvector Network Centrality measures the informational advantage a bond fund has based on the number of connections it has and the importance of bond funds it links to. Eigenvector Network Centrality is standardized in regression. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Fund fixed effects are included. Stars represent statistical significance. (***p<0.001, **p<0.005, *p<0.1, respectively)

	Network Centrality		
	(1)	(2)	(3)
Low Ranked Fund	0.031**		0.035***
	(0.013)		(0.013)
High Ranked Fund		0.011	0.017
		(0.012)	(0.012)
Total Net Assets	-0.040***	-0.040***	-0.040***
	(0.008)	(0.008)	(0.008)
Trading Volume	0.009	0.009	0.009
	(0.008)	(0.008)	(0.008)
Traded Bond Issue Counts	0.097***	0.096***	0.097***
	(0.015)	(0.015)	(0.015)
Fund Fixed Effect	Yes	Yes	Yes
Observations	20,632	20,632	20,632
R^2	0.019	0.018	0.019

Table 1.8: Within Fund Time Varying Changes of Network Centrality: Past Fund Flow

This table presents the within-find time varying change of eigenvector network centrality due to past fund flow fluctuations. Past Fund flow is net fund flow each quarter. Eigenvector Network Centrality measures the informational advantage a bond fund has based on the number of connections it has and the importance of bond funds it links to. Eigenvector Network Centrality is standardized in regression. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Fund fixed effects are included. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Network Centrality	
	t+1	t+2
Past Fund Flow	0.281***	0.162***
	(0.040)	(0.042)
Total Net Assets	-0.023***	-0.026***
	(0.008)	(0.009)
Trading Volume	-0.005	-0.002
	(0.008)	(0.008)
Traded Bond Issue Counts	0.049***	0.026
	(0.016)	(0.017)
Fund Fixed Effects	Yes	Yes
Observations	19,281	18,522
R^2	0.006	0.002

Table 1.9: Network Centrality-Fund Performance Relationship: Leader and Follower

Following table in this section present the directed information network of herding behavior with herding in a one quarter lag and herding with two quarters lag respectively. I classify bond mutual funds in the top quantile based on their leading legs in the directed information network. Leader funds are funds in the top quantile which have the most number of other funds following their investment decisions. Follower funds are funds in the bottom quantile which mainly follow other funds in making an investment decision. The opposite regression coefficient on Leader vs Follower variables indicates that network central funds capture profitable information. Different from previous networks focusing on the same trades within the same quarter, herding information network is constructed based on the same trades sharing by bond funds across time periods. In each quarter, a herding link from A fund to B fund is formed if A executes the same trades which B bond fund already transacted one (two) quarter ago. These networks allow us to study the herding behavior among bond funds. Alpha is future risk adjusted fund performance, which is the annualized version of monthly alpha. Monthly alpha is calculated by running 12-month rolling regression of bond fund excess net return on the aggregate stock market excess return from CRSP and the aggregate Vanguard total bond market index fund excess return, following Goldstein, Jiang, and Ng (2017). 1-month, 3-month and 12-month are bond fund future net return up to one year. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of a number of traded bond issues by each fund quarterly. Quarterly Fund Flow is the percentage of fund low in a given quarter. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Annualized Alpha	
	One Period	Two Periods
Leader	1.207***	1.232***
	(0.290)	(0.286)
Follower	-0.921***	-0.796***
	(0.292)	(0.282)
Total Net Assets	0.072	0.071
	(0.090)	(0.090)
Trading Volume	-0.333***	-0.338***
	(0.107)	(0.107)
Traded Bond Issue Counts	0.459**	0.503**
	(0.221)	(0.218)
Observations	20,777	20,777
R^2	0.002	0.002

Table 1.10: Network Centrality-Fund Performance Relationship During the Periods of High Market Uncertainty and For Bond Funds with High Fund Level Liquidity

This table presents how the relationship between bond mutual funds network centrality and fund future returns change both on time series perceptive (during the periods of uncertainty) and cross-sectional perceptive (funds with different level of holding liquidity). The underlying information network is constructed with nodes representing each bond mutual funds and linkage between any nodes is formed if any two bond funds have the same trades (either sell or buy) during the same quarter. The weight of the edge is equal to the percentage of overlapping trades they share out of their corresponding quarterly trading quantities. The weight measures the intensity of overlapping trading behavior between any two mutual funds in a govern quarter. Eigenvector Network Centrality measures the informational advantage a bond fund has based on the number of connections it has and the importance of bond funds it links to. Eigenvector Network Centrality is standardized in regression. 12-month Alpha is future risk adjusted fund performance, an annualized version of monthly alpha. Monthly alpha is calculated by running 12-month rolling regression of bond fund excess net return on the aggregate stock market excess return from CRSP and the aggregate Vanguard total bond market index fund excess return, following Goldstein, Jiang, and Ng (2017). 1-month, 3-month and 12-month are bond fund future net return up to one year. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Crisis is the dummy variable for the crisis period (2007Q3 to 2009Q2). Eigenvector Centrality*Crisis is the interaction terms between Eigenvector Network Centrality and Crisis. Financial Stress Index is Financial Stress Index St. Lois Fed which captures the degree of financial stress and market uncertainty in the financial markets. Eigenvector Centrality*Financial Stress Index is the interaction term. Both Crisis dummy and Financial Stress Index are proxies for the period of market uncertainty. Fund Holdings liquidity is the fund level liquidity. It is calculated by aggregating the Amihud illiquidity measure (Amihud (2002)) of each compote bond in fund holdings with holdings as weight. I multiply this aggregated number by -1 to show fund level liquidity in the regression. Eigenvector Centrality*Fund Holdings Liquidity is an interaction term. Cash Equivalents Holdings is the percentage of cash and government bonds held by funds in a given quarter. Eigenvector Centrality*Cash Equivalents Holdings is the interaction term. Both Fund Holdings liquidity and Cash Equivalents Holdings are proxies for fund level liquidity. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Aggregate Information Network			
	Annualized Alpha			
	(1)	(2)	(3)	(4)
Eigenvector Centrality	0.004**	0.008***	0.003*	0.003
	(0.002)	(0.002)	(0.001)	(0.002)
Total Net Assets	-0.010***	-0.003	0.002**	0.003***
	(0.002)	(0.002)	(0.001)	(0.001)
Trading Volume	0.001	-0.002	-0.004***	-0.005***
	(0.002)	(0.002)	(0.001)	(0.001)
Traded Bond Issue Counts	-0.007**	0.001	0.010***	0.009***
	(0.003)	(0.003)	(0.002)	(0.002)
Crisis	0.023***			
	(0.003)			
Eigenvector Centrality*Crisis	0.025***			
	(0.003)			
Financial Stress Index		0.019***		
		(0.001)		
Eigenvector Centrality*Financial Stress Index		0.012***		
		(0.001)		
Fund Holdings Liquidity			0.003**	
			(0.001)	
Eigenvector Centrality* Fund Holdings Liquidity			0.006***	
			(0.001)	
Cash Equivalents Holdings				0.0002
				(0.0003)
Eigenvector Centrality* Cash Equivalents Holdings				0.001**
				(0.0004)
Fixed Effects	Yes	Yes	Yes	Yes
Observations	20,777	20,777	18,256	20,321
R^2	0.015	0.038	0.003	0.003

Figure 1.1: Total Net Value of Corporate Bond Mutual Funds

This figure presents the total net assets of corporate bond mutual funds in the U.S. It includes bond mutual funds with investment objective in both investment grade bonds and high yield bonds. The data source is from Investment Company Institution.

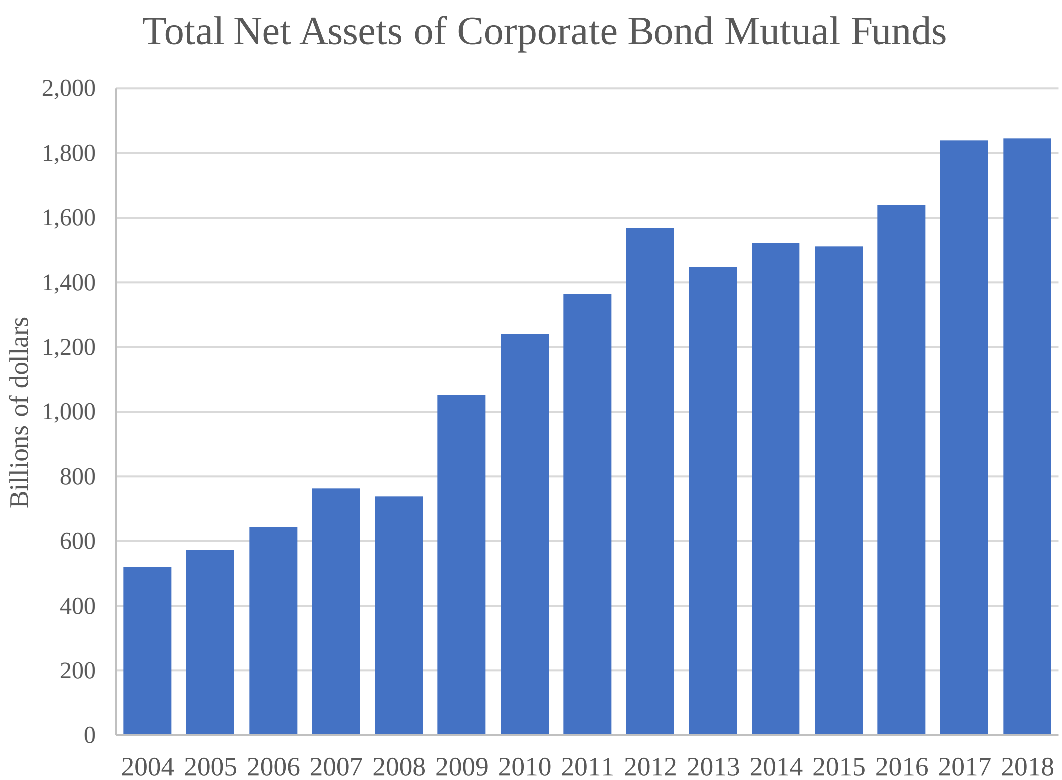
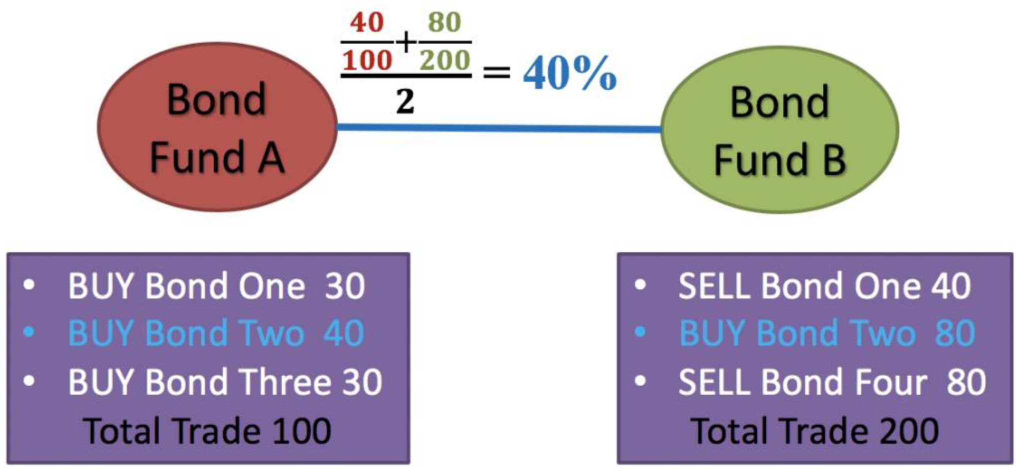


Figure 1.2: Weighted Information Network Construction

This figure exhibits the construction of weighted information network of corporate bond mutual funds. Bond Fund A and Bond Fund B are connected by their common transaction on Bond Two at the same time with the same buying direction. The linkage of these two bond funds is weighted by their scaled overlapping trading volume.



Chapter 2 : Information Networks and Corporate Bonds

2.1 Introduction

This chapter examines the impact of the information network of corporate bond mutual funds on their underlying corporate bonds. Chapter 1 constructs the information networks of bond mutual funds according to their quarterly trading behavior to study the information flow between corporate bond mutual funds. With a specific focus on the impacts of information networks on fund performance and fund herding behaviors, chapter 1 shows that corporate bond funds placed in the central portion of the information networks lead to higher fund future returns due to their informational advantages, and intensified herding behavior from other bond funds. A natural question that follows is how information networks of bond mutual funds influence underlying corporate bonds, besides its impact on fund level analysis. Does trading by corporate bond funds improve underlying corporate bond liquidity and increase the information embedded in prices? In chapter 2, the main focus is upon corporate bond level analysis.

I first investigate the effect of network centrality on the underlying corporate bond liquidity. I find that corporate bonds owned by highly network-central funds increase underlying bond liquidity. To reach this conclusion, I construct a network centrality weighted fund ownership for each bond. One standard deviation increase in bond ownership by high network centrality funds is found to lead to a 2.2% decrease in Amihud illiquidity measure and a 3.5% decrease in the average bid-ask spread next quarter. This significant positive relationship between network centrality and corporate bond liquidity persists through various different measurements of bond liquidity. Moreover, given that each liquidity measurement only captures the portion of information content in bond liquidity, I adopt the principal component method used by Korajczyk and Sadka (2008) and Dannhauser (2017) to construct an aggregate measurement of bond liquidity. The bond ownership by high network centrality fund again leads to an increase in aggregate liquidity measure. Hence, it establishes a robust positive relationship between network centrality and bond liquidity. In addition, I also study the impact of network centrality on other bond level characteristics including bond return and bond trading volume.

After showing this positive relationship between network centrality and bond liquidity, I further study the mechanism behind this positive relationship. The positive relationship is hypothesized to occur due to increased information sharing through the network which intensifies herding behavior. Trading from bond funds with high network centrality improves pricing efficiency by enabling more bond specific information to be reflected in the bond price. The influence of high centrality funds causes other bond funds to trade on the same bond as shown by increments in future trading volume and herding intensities. By adopting the 1-R square method by Chen, Goldstein, and Jiang (2007), I find that one standard deviation increase in bond ownership by high network centrality funds leads to 9.3% increase in bond pricing informativeness. Improved information efficiency and herding behaviors lead to increases in bond liquidity.

This chapter makes three important contributions. First, it uncovers the important differential force on corporate bond liquidity exerted by corporate bond mutual funds according to their different positions in the information networks. Second, it uniquely identifies trading information sharing among corporate bond mutual funds that are consistent across quarters from their actual trades. It presents the first empirical evidence on herding behaviors which directly examines each leader and follower pair on every trade. Third, it highlights the improved pricing efficiency in the corporate bond market brought about by trading behaviors of highly-connected information network bond funds.

The reminder of the paper is organized as follows. Section 2.2 presents related literature and contribution. Section 2.3 shows main hypotheses development. Section 2.4 discusses the empirical results of the impact of information networks of corporate bond mutual funds on their underlying corporate bond market. Section 2.5 concludes.

2.2 Contribution to Literature

This chapter contributes to two strands of literature. First, it contributes to the growing literature on corporate bond market liquidity. It still remains a topic of debate as to whether corporate bond market liquidity decreased or increased after dealers started retreating from market marking activities (Harris (2015), Adrian, Flemming, Shachar, Vogt (2016) and Choi and Huh (2017)). As far as I know, this work presents the first evidence that holdings by bond funds with high network centrality improves underlying bond liquidity. Trading by informative network-central bond funds results in more bond specific information to be included in prices and more herding behavior on the same bond. My evidence highlights the differential influence on bond liquidity by corporate bond funds in respect to their different network positions.

Second, my work also contributes to the large literature studying herding behavior of institutional investors in the capital markets. Most of the current works focus on the equity market. bond market. Prior papers have studied the price impacts of institutional herding in stocks. Dasgupta, Prat, and Verardo (2011), Brown, Wei, and Wermers (2013), and Singh (2013) shows price reversals after herding, whereas Sias (2004) do not find the evidence of stock price reversal after herding. Cai, Song, Dan and Yi (2019) are the first to study the herding behavior of institutional investors in the corporate bond market. They find that institutional herding destabilizes corporate bond prices. Cai, Song, Dan and Yi (2019) identify herding intensity by comparing the selling and buying volume of corporate bond with the average trading volume of all the bonds, following the methods done in the equity literature (Brown, Wei, and Wermers (2013)). My network analysis builds upon such work by being the first to directly examining each leader and follower pair on every trade and construct time series fund level herding intensity.

2.3 Main Hypothesis Development

Different from equity markets, corporate bonds trade over the counter. The opacity of corporate bond trading causes corporate bonds to have less information efficiency compared to its counterparty. Harris (2015), O'Hara, Wang, and Zhou (2016) and Babus and Kondor (2018) have shown that high searching cost and the loss of information resulting from the lack of a centralized corporate bond trading platform. These frictions amplify the importance of information in the corporate bond market. The institutional background of the corporate bond market leads me to the following three hypotheses in this section.

Hypothesis I: Bonds in the portfolio of high centrality funds have higher future liquidity

Bonds owned and traded by the highly connected and network central bond funds leads to higher future corporate bond liquidity, due to two channels: intensified institutional herding and improved information efficiency. The trading behavior by informed network-central bond funds increases the bond specific information reflected in the price. It further leads to more bond funds mimicking the investment decision. Hence, it results in overall increased trading volume and liquidity of the same bond.

Hypothesis II: Increased liquidity results from intensified herding behavior

Corporate bond mutual funds in the periphery of the information network tend to follow information leader, network central corporate bond funds, by trading on the same corporate bonds. The increased trading volume due to intensified herding behavior contributes to the increase in the underlying corporate bond liquidity.

Hypothesis III: Bonds traded by network central bond funds enhances bond pricing efficiency, with more corporate bond specific information to be incorporated in the price.

Due to informational advantage by being in the center of the information network, network central bond mutual funds are more likely to make the informed trading decision. Their trading activities help incorporate fundamental information in the underlying corporate bond prices.

2.4 Data Construction

Sample data on corporate bonds is primiarly constructed from following three databases: Thomson Reuters Lipper eMaxx database, Fixed Investment Securities Database, Enhanced Historic Trade Reporting and Compliance engine.

2.4.1 Corporate bond pricing and trading related data

TRACE Enhance database provides comprehensive over-the-counter corporate bond transaction data starting from 2002 to 2014. After retrieving complete historical transaction data from TRACE, I adopt methods described in Dick-Nielsen, Feldhütter, and Lando (2012) and Dick-Nielsen (2014) to clean the TRACE database to correct for various erroneous trades, liquidity biases, cancellation, and double counted trades. Next, I merge the TRACE Enhance database with Fixed Investment Securities Database by security level eight-digit CUSIP to obtain detailed corporate bond panel data regarding bond pricing and trading. Fixed Investment Securities Database offers detailed bond issuer and bond issue level characteristics (bond ratings and time to maturity, etc.). To obtain bond rating each quarter, I apply the highest rating of the major credit rating agencies (S&P, Moody's and Fitch). Armed with high-frequency transaction level data, I am able to construct various liquidity measurements and corporate bond return measurement.

2.4.2 Corporate bond holding data

Thomson Reuters eMaxx database contains the most comprehensive corporate bond security level information regarding bond holdings by major institutional investors. It provides end of quarter holding positions by mutual funds, pension funds, and insurance companies. eMaxx database is a superior source of corporate bond holding data as it includes all defunct fund shops and is free of reporting biases and survivorship biases. I first retrieve the holdings information about the corporate bond mutual funds in the U.S. from eMaxx database. I then follow previous literature to make the dataset ready for use (Massa and Zhang (2012), Becker and Ivashina (2015), Choi and Kronlund (2017) and Cai, Song, Dan and Yi (2019)). Lastly, I merge the corporate bond holding positions with the detailed corporate bond pricing data I obtain in the previous section to create the detailed corporate bond panel data used in this study.

2.4.3 Liquidity variables construction

I construct various corporate bond liquidity variables that are commonly used in the prior literature. Amihud (2002)'s price impact measurement which calculates the ratio of the price movement over total dollar trading volume. It is calculated as

$$BondLiquidity_{i,D} = \frac{1}{N_{i,D}} \sum\nolimits_{k} \frac{\frac{|Price_{i,D}^{k} - Price_{i,D}^{k-1}|}{Price_{i,D}^{k-1}}}{TradeQuantity_{i,D}^{k}}$$

where $Price_{i,D}^{k}$ is the price of *k*-th trade on the bond *i* on day *D* during the quarter t. $N_{i,D}$ is number of trades on bond *i* on day D.

I also construct Lesmond, Ogden, and Trzcinka (1999)'s imputed roundtrip cost bond liquidity measure. It calculates the transaction cost through bond trades executed within a short window of the similar trading volume.

$$ImputedRC_{\mathrm{i,t}} = \frac{Price_{max,i,t} - Price_{min,i,t}}{Price_{min,i,t}} * 100$$

Lesmond, Ogden, and Trzcinka (1999)'s zero trading days is the other liquidity measurement. It measures the stagnancy of corporate bonds which have not been traded for a long period. The longer duration the zero trading days is, the less liquid a bond is.

$$ZeroTradingDays_{i,t} = \frac{\text{number of zero return days}_{i,t} + number\ of\ zero\ trade\ days_{i,t}}{total\ number\ of\ trade\ days_t} * 100$$

Dick-Nielsen, Feldhütter, and Lando (2012)'s interquartile price range is another commonly used liquidity measurement. It is measured as the difference between the interquartile prices of trade bonds. Lastly, the average trade spread is calculated to measure bond level liquidity. It is the average of quarterly bid-ask spreads for bonds.

2.5 Empirical Analysis: Impact of Information Networks on Corporate Bonds

In this section, I present the empirical analysis of the influence of the information network of bond funds on their underlying corporate bond. I start with the panel fixed effect regression model to show that corporate bonds traded by network central bond funds lead to an increase in the underlying corporate bond liquidity. I further confirm this result with various commonly used liquidity measurement in the academic papers, in addition to an aggregate liquidity measurement generated from the principal component method. After establishing the positive relationship between network centrality and the underlying corporate bond liquidity, I examine two channels via which potentially forms the relationship. The first channel is the intensified herding behaviors from other bond funds following the network central bond funds to trade on the same bonds. It associates with intensified herding behavior and a boost in the trading volume. The other channel is that trading by informed network central bond funds which enables more bond specific information to be reflected in the bond price. As a result, the improved bond pricing efficiency contributes to the increase in the underlying bond liquidity.

2.5.1 Increase in holding by network central funds improve bond future liquidity

To examine the effect of the ownership by network central bond funds on the liquidity of the underlying corporate bonds, I conduct the following panel fixed effect regression.

$$Liq_{i,t+1} = \beta_1 OwnershipWeightedbyCentrality_{i,t} + Controls_t + FixedEffects + \epsilon_{i,t}$$

$$OwnershipWeightedbyCentrality_{i,t} = \sum_j \omega_{i,j,t} NetworkCentrality_{j,t}$$

$Liq_{i,t+1}$ is corporate bond liquidity. $OwnershipWeightedbyCentrality_{i,t}$ represents the aggregated ownership of corporate bond *i* weighted by the eigenvector network centrality of bond mutual funds during quarter t. $\omega_{i,j,t}$ denotes the ownership of corporate bond *i* by bond mutual fund *j* at time *t*. The higher this measure indicates a larger proportion of underlying corporate bond held by more connected network central bond funds.

[Insert Table 2.1 here]

I adopt five bond liquidity measures from the literature to show the robustness of the result. Column 1 to 5 in Table 2.1 show that one standard deviation increased in the bond ownership by high network central funds lead to a decrease in future bond Amihud illiquidity by 2.2%, a decrease in zero trading days next quarter by 1%, a decrease in interquartile price range next quarter by 0.2 %, a decrease imputed round-trip cost by 2%, and a decrease in average bid-ask spread next quarter by 3.5%. The bond ownership by network central bond funds universally leads to a decrease in bond level future illiquidity (thus, a boost in the bond liquidity), regardless of illiquidity measurements chosen.

[Insert Table 2.2 here]

Each liquidity measurement captures a portion of the information contained in the bond liquidity. Hence, to further confirm this positive relationship between bond fund network centrality and underlying bond liquidity, I construct an aggregate liquidity proxy using principal component analysis (PCA) methods, used in Korajczyk and Sadka (2008) and Dannhauser (2017). Principle component analysis enables me to construct a superior liquidity measurement which captures the aggregate information from various proxies for bond liquidity. Following closely Korajczyk and Sadka (2008), I compute the first principal component of the PCA method from 5 liquidity measures mentioned above as the aggregate bond liquidity measurement. From the estimated coefficient, I confirm that an increase in the bond ownership by network central bond funds leads to the increased liquidity of their underlying corporate bonds.

2.5.2 Increase in holdings by network central funds comes with increase in bond trading volume

$$M_{i,t+1} = \beta_1 OwnershipWeightedbyCentrality_{i,t} + Controls_t$$
$$+FixedEffects + \epsilon_{i,t}$$

[Insert Table 2.3 here]

Moreover, I investigate the impact of trading behavior of network central funds on other bond characters like bond future return and bond trading volume. $M_{i,t+1}$ in above regression denotes bond level characteristics. Table 2.3 shows that one standard deviation increase in the bond ownership by high network central funds increase bond trading volume in the next quarter by a 0.1 standard deviation, and increase future bond return by 0.2%. The above mentioned the increased in bond level liquidity comes with this surge in bond trading volume. These observations lead me to propose the following two potential channels that explain the increased bond level liquidity by bond ownership of network central bond funds.

[Insert Table 2.4 here]

2.5.3 Increased bond future liquidity comes from channel I: intensified herding behavior

To measure the herding behavior directly, I construct a herding information network. The herding network is built using the same bond trades following across quarters, rather than using the same trades executed within the same quarter in the networks as shown previously. For each trade placed during a given quarter t, I track the trades on the same bonds by other bond funds occurring during the quarter t+1 and quarter t+2. By doing so, I am able to examine the herding behavior over time more granularly at a leader-follower trade level, instead of at an aggregate level as commonly done in the prior literature. I find that one standard deviation increase in bond fund eigenvector centrality leads to a 0.2 standard deviation increase in herding intensity in the following quarter. This lends support to my hypothesis by presenting a direct evidence of increased herding behaviors. I further report in Table 2.5 the impact of herding across time periods on bond fund return. One standard deviation increase in one bond fund's herding behaviors could lead to a 3% decrease in bond fund's future risk adjusted return.

[Insert Table 2.5 here]

The increase in liquidity could come from herding behaviors by other bond funds in the periphery of the network. Trading by bond funds with high centrality increase pricing efficiency by enabling more bond specific information to be incorporated into the price. Other bond funds in the network tend to herd the high network central bond funds and trade on the same bonds. It thus leads to an increase in trading next quarter and increases in bond return. I show in the previous section that a 0.1 standard deviation increase in next quarter trading volume associated with one standard deviation increase in the ownership of bonds owned by high network central bond funds.

2.5.4 Increased bond future liquidity comes from channel II: improved bond price informativeness resulting from trading by high network central funds.

$$BondPriceInformativeness_{i,t+1} = \beta_1 OwnershipWeightedbyCentrality_{i,t}$$
$$+Controls_t + FixedEffects + \epsilon_{i,t}$$

[Insert Table 2.5 here]

I adopt 1-R squared method to compute the bond price Informativeness, following the method used in Chen, Goldstein, and Jiang (2007). From Table 2.5, I find that one standard deviation increases in ownership by network central bond funds leads to a 9% increase in bond pricing efficiency. It indicates that trading by network central bond funds makes more bond specific information to be incorporated into bond pricing. This increased in bond pricing efficiency contributes to the increase in bond liquidity.

2.6 Conclusion

This chapter studies the impacts of information networks of corporate bond mutual funds on the underlying corporate bonds. I show that an increase in bond ownership by bond funds positioning in the central portion of the networks leads to increase in their underlying corporate bond liquidity. The positive relationship between network centrality and bond liquidity is robust for various bond liquidity measurements. Moreover, I find that the increased liquidity results from the intensified herding behavior of bond funds following network central bond funds by tading on the same bonds, and improved bond pricing informativeness. This study provides the first empirical evidence shows that the differential influence of corporate bond mutual funds on their underlying bonds liquidity due to their different positions in the information network.

2.7 Figures and Tables

Table 2.1: Effect of Network Centrality of Bond funds on Underlying Corporate Bond Liquidity

This table presents how the trading and holding by high network-central corporate bond mutual fund affects the underlying cooperate bond liquidity. Eigenvector Centrality Weighted Bond Ownership is constructed at the bond level in each given quarter. For every corporate bond ever traded by any corporate bond funds, I multiply the percentage of the bond held by each corporate bond fund with its corresponding bond fund network centrality and sum this number across all the bond funds who hold this bond in the given quarter. This variable measures the intensity of a corporate bond owned by high network centrality bond fund. Bond Rating is the max of ratings coming from three credit rating agencies: Moody's, S&P and Fitch. Maturity is time to maturity. I use five liquidity proxies. Amihud Illiquidity is the bond level illiquidity measurements (Amihud (2002)). Zero Trading days measure no trading days or zero return over total trading days (Lesmond, Ogden, and Trzcinka (1999)). Interquartile Price Range is the interquartile price difference of traded bonds (Dick-Nielsen et al (2012)). Imputed Roundtrip Cost is the measurement of transaction cost constructed following (Feldhütter (2012)). Average Trade Spread is the quarterly average bid-ask price for each bond. Bond and time fixed effects are included. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Amihud Illiquidity	Zero Trading days	Interquartile Price Range	Imputed Roundtrip Cost	Average Trade Spread
	(1)	(2)	(3)	(4)	(5)
Eigenvector Centrality Weighted Ownership	-0.022***	-0.010***	-0.0002***	-0.020***	-0.035***
	(0.001)	(0.001)	(0.00002)	(0.002)	(0.003)
Bond Rating	0.006***	-0.011***	0.0002***	0.007***	0.016***
	(0.001)	(0.0005)	(0.00001)	(0.001)	(0.002)
Maturity	0.001***	-0.004***	0.00002***	0.003***	0.001*
	(0.0001)	(0.0001)	(0.00000)	(0.0002)	(0.0003)
Bond Fixed Effects	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	88,531	90,304	90,304	82,199	85,154
R^2	0.126	0.221	0.197	0.213	0.254

Table 2.2: Effect of Network Centrality of Bond funds on Corporate Bond Liquidity Using PCA Method

Table 2.2 presents the effect of corporate bond level network-centrality measure underlying cooperate bond liquidity using principal component analysis (PCA). The dependent variable representing the bond level liquidity measurement which is the first principal component of the PCA method. Principle component analysis is able to capture all the information from various liquidity measures using in the past literature (Amihud Illiquidity, Zero Trading days, Interquartile Price Range, Imputed Roundtrip Cost, and Average Trade spread). Principle component analysis is conducted following Korajczyk and Sadka (2008). Eigenvector Centrality Weighted Bond Ownership is constructed at the bond level in each given quarter. For every corporate bond ever traded by any corporate bond funds, I multiply the percentage of the bond held by each corporate bond fund with its corresponding bond fund network centrality and sum this number across all the bond funds who hold this bond in the given quarter. This variable measures the intensity of a corporate bond owned by high network centrality bond fund. Bond Rating is the max of ratings coming from three credit rating agencies: Moody's, S&P and Fitch. Maturity is time to maturity. Bond and time fixed effects are included. Stars represent statistical significance. (***p<0.001, **p<0.005, *p<0.1, respectively)

	Liquidity Measurement
	First Principal Component of PCA Method
Eigenvector Centrality Weighted Ownership	0.115***
	(0.010)
Bond Rating	-0.049***
	(0.006)
Maturity	-0.018***
	(0.001)
Bond Fixed Effects	Yes
Time Fixed Effects	Yes
Observations	76,726
R^2	0.286

Table 2.3: Effect of Network Centrality of Bond funds on Underlying Corporate Bond Market

Table 2.3 presents how the trading and holding by high network-central corporate bond mutual fund affects the underlying cooperate bond return and trading volume. Eigenvector Centrality Weighted Bond Ownership is constructed at the bond level in each given quarter. For every corporate bond ever traded by any corporate bond funds, I multiply the percentage of the bond held by each corporate bond fund with its corresponding bond fund network centrality and sum this number across all the bond funds who hold this bond in the given quarter. This variable measures the intensity of a corporate bond owned by high network centrality bond fund. Bond Rating is the max of ratings coming from three credit rating agencies: Moody's, S&P and Fitch. Maturity is time to maturity. Bond Return is quarterly corporate bond return. Trading Volume is the quarterly trading volume. Bond and time fixed effects are included. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Bond Return	Trading Volume
	(1)	(2)
Eigenvector Centrality Weighted Ownership	0.002***	0.102***
	(0.001)	(0.013)
Bond Rating	0.003***	0.030***
	(0.0005)	(0.008)
Maturity	0.0004***	0.017***
	(0.00005)	(0.001)
Bond Fixed Effects	Yes	Yes
Time Fixed Effects	Yes	Yes
Observations	87,868	86,349
R^2	0.067	0.098

Table 2.4: First Channel of Increased Bond Liquidity and Trading Volume: Intensified Herding Behaviors

Table 2.4 presents the first possible channel of increased underlying bond level liquidity and trading volume results from intensified herding behaviors of other funds brought by the trading of high network central funds. The increasing herding behaviors from other funds following high network central funds trading on the same bonds. The dependent variable is herding intensity measured by the number of followers in the following quarter. Column 1 presents the standardized herding intensity and column 2 is the raw number. Eigenvector Network Centrality is standardized in regression. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of the number of traded bond issues by each fund quarterly. Bond and time fixed effects are included. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Herding Intensity	
	(1)	(2)
Eigenvector Centrality	0.184***	13.241***
	(0.009)	(0.633)
Total Net Assets	0.023**	1.675**
	(0.009)	(0.678)
Trading Volume	-0.019**	-1.339**
	(0.008)	(0.603)
Traded Bond Issue Counts	0.510***	36.752***
	(0.021)	(1.487)
Bond Fixed Effect	Yes	Yes
Time Fixed Effect	Yes	Yes
Observations	20,277	20,277
R^2	0.597	0.597

Table 2.5: Intensified Herding Behaviors Lead to Decreases in Herders' Fund Alpha

Following two tables in this section present the information network of herding behavior with herding in a one-quarter lag and herding with two quarters lag respectively. Different from previous networks focusing on the same trades within the same quarter, herding information network is constructed based on the same trades sharing by bond funds across time periods. In each quarter, a herding link from A fund to B fund is formed if A executes the same trades which B bond fund already transacted one (two) quarter ago. These networks allow us to study the herding behavior among bond funds. Eigenvector Network Centrality in these two herding networks measures the intensity of herding behavior. The funds in the center of herding network herd most. I can see that funds herd most with two-quarter lag results in a significant decrease in future fund performance. In the regression, Eigenvector Network Centrality is in a standardized form. 12-month Alpha is future risk adjusted fund performance, which is the annualized version of monthly alpha. Monthly alpha is calculated by running 12-month rolling regression of bond fund excess net return on the aggregate stock market excess return from CRSP and the aggregate Vanguard total bond market index fund excess return, following Goldstein, Jiang, and Ng (2017). 1-month, 3-month and 12-month are bond fund future net return up to one year. Log Total Net Assets is the logarithm of bond fund total net assets. Log Trading Volume is the logarithm of the total quarterly trading quantity by each fund. Log Traded Bond Issues is the logarithm of a number of traded bond issues by each fund quarterly. Quarterly Fund Flow is the percentage of fund low in a given quarter. Fund fixed effects are included in the regression. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Information Network on Herding behavior (herd in one quarter)			
	1-month Return	3-Month Return	12-month Return	Annualized Alpha
	(1)	(2)	(3)	(4)
Eigenvector Network Centrality	0.001***	0.003***	0.005***	-0.001
	(0.0003)	(0.0005)	(0.001)	(0.002)
Total Net Assets	0.0001	-0.003***	-0.015***	-0.012***
	(0.0003)	(0.0004)	(0.001)	(0.002)
Trading Volume	0.0004*	0.002***	0.007***	0.001
	(0.0003)	(0.0004)	(0.001)	(0.002)
Traded Bond Issue Counts	0.001	-0.002***	-0.015***	-0.007*
	(0.001)	(0.001)	(0.002)	(0.003)
Fund Fixed Effect	Yes	Yes	Yes	Yes
Observations	20,246	20,261	20,261	20,261
R^2	0.004	0.005	0.014	0.005

	Information Network on Herding behavior (herd in two quarters)			
	1-month Return	3-Month Return	12-month Return	Annualized Alpha
	(1)	(2)	(3)	(4)
Eigenvector Network Centrality	0.001***	0.0004	-0.005***	-0.003***
	(0.0003)	(0.0004)	(0.001)	(0.001)
Total Net Assets	0.0002	-0.003***	-0.015***	-0.010***
	(0.0003)	(0.0004)	(0.001)	(0.001)
Trading Volume	0.0004*	0.002***	0.007***	0.001
	(0.0003)	(0.0004)	(0.001)	(0.001)
Traded Bond Issue Counts	0.0005	-0.002**	-0.013***	-0.006***
	(0.001)	(0.001)	(0.002)	(0.002)
Fund Fixed Effect	Yes	Yes	Yes	Yes
Observations	19,735	19,750	19,750	19,750
R^2	0.005	0.003	0.012	0.011

Table 2.6: Second Channel of Increased Underlying Liquidity: Improved Pricing Efficiency Brought by Holding and Trading Behaviors of Network Central Funds: R squared Method

Table 2.6 presents the improved pricing efficiency brought by holding and trading behavios of network central funds. Eigenvector Centrality Weighted Bond Ownership is constructed at the bond level in each given quarter. For every corporate bond ever traded by any corporate bond funds, I multiply the percentage of the bond held by each corporate bond fund with its corresponding bond fund network centrality and sum this number across all the bond funds who hold this bond in the given quarter. This variable measures the intensity of a corporate bond owned by high network centrality bond fund. Bond Rating is the max of ratings coming from three credit rating agencies: Moody's, S&P and Fitch. Maturity is time to maturity. Bond pricing informativeness is constructed using R squared method adopted in Qi, Goldstein, Jiang (2010). Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Bond Pricing Informativeness
	(1)
Eigenvector Centrality Weighted Ownership	0.993***
	(0.186)
Bond Rating	0.182*
	(0.097)
Maturity	-0.108***
	(0.018)
Bond Fixed Effects	Yes
Time Fixed Effects	Yes
Observations	89,367
R^2	0.096

Figure 2.1: The U.S. Corporate Bond Issuance Market

This figure exhibits the total size of the corporate bond Issuance Market in the U.S. It shows the rapid growth of the size of corporate bond issuance. The data source is obtained from the Securities Industry and Financial Markets Association (SIFMA).

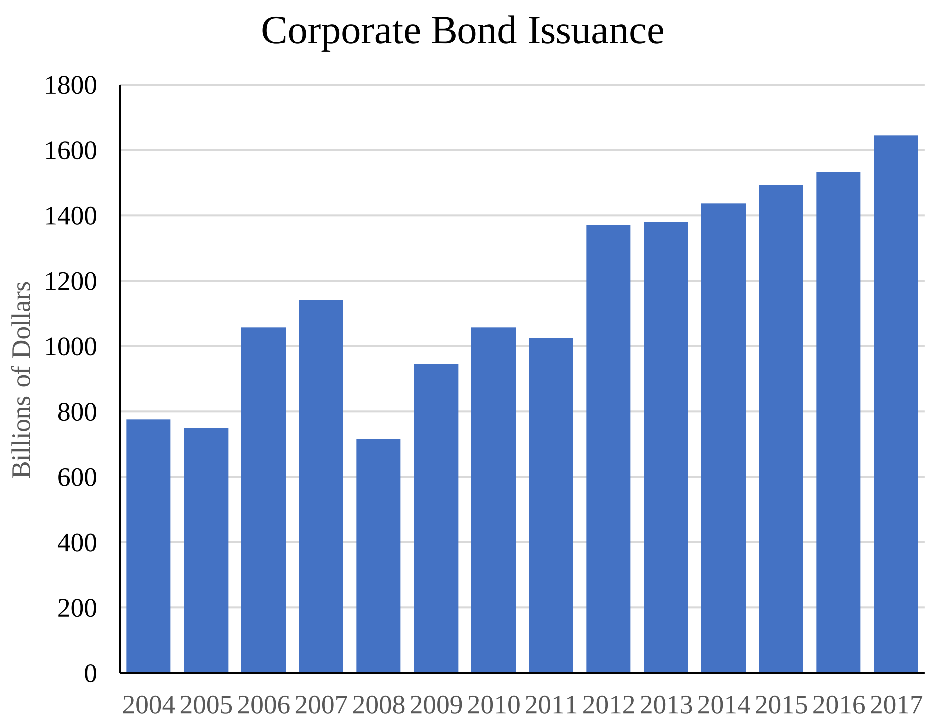
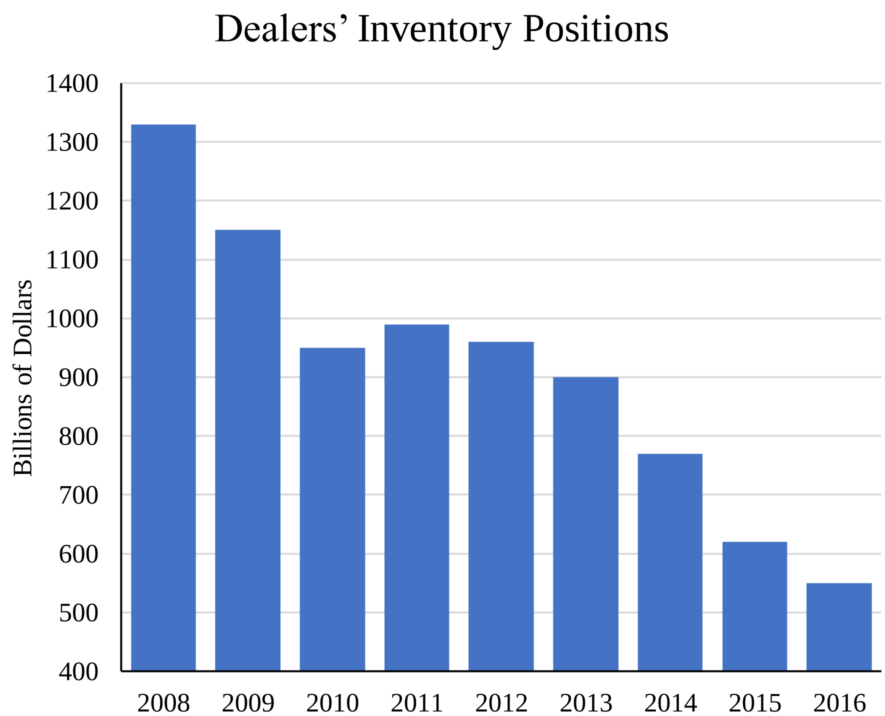


Figure 2.2: Dealers' Inventory Positions

This figure exhibits primary dealers' inventory positions in the bond market since the financial crisis. It shows a huge decline in dealers' inventory positions as dealers retreat from market making. The data is obtained from the Federal Reserve Board, Form FR 2004A.



Chapter 3 : Corporate Bond Exchange-traded Fund and Corporate Bond Return Comovement

3.1 Introduction

The corporate bond market is an important place for firms to raise capital, with the value of corporate bond issuance market reaching a record $1.65 trillion (SIFMA 2018). The corporate bond market used to rely on broker dealers to provide liquidity. However, despite record growth in the issuance market, there has been more than 70% decline in dealer's inventory position (Federal Reserve Board 2016). As dealers retreat from market making, other non-bank participants are increasingly an important role in this market. Under this background, corporate bond exchanged-traded fund (corporate bond ETF) has seen an over 3000% growth in asset under management over the past 10 years (Investment Company Institute 2017). Corporate bond ETFs have become one of the fastest growing assets in the fixed income market. The growing popularity of corporate bond ETF makes bond indexing more accessible and cheaper than ever before. As corporate bond ETF become increasingly more popular, it is important to understand how corporate bond ETFs affect underlying corporate bond price movement and how it differs from other traditional institutions.

There are growing concerns among the public regarding the pricing impacts of ETFs on the underlying securities. In a frictionless market, basket securities such as bond ETFs are redundant assets that have no impacts on the price movement of the underlying securities. However, bond ETFs are invested in the underlying corporate bonds using a mechanical approach according to their weights in the indexes tracked by bond ETFs (Dannhauser (2017)). Demand for corporate bonds resulting from the mechanical part of the ETF might become disconnected from the fundamental information for which individual bonds are exposed to. For example, if a bond faces declining demand due to negative information, capital flows from corporate bond ETFs will still push up demand for this bond and cause the price to move away from its fundamental value. As a result, ETFs transmit non-fundamental shocks to the underlying corporate bonds. This makes the prices of corporate bonds that are part of ETF portfolios move in synchrony more than what they are supposed to, purely based on fundamental information. This paper fills the gap in the literature by providing the first empirical evidence on whether and how corporate bond ETFs impact corporate bond price comovement.

My study begins by examining if corporate bond ETFs ownership makes corporate bonds returns comove more with other corporate bonds in the same bond ETFs. I adopt the method used in Da and Shive (2018) to measure the intensity of return comovement. Since a corporate bond is held by multiple ETFs, a super portfolio which includes corporate bonds owned by these ETFs is constructed. The intensity of bond return comovement is measured by the beta of a corporate bond return's with respect to the return of the super portfolio and the return of the aggregate corporate bond market. If corporate bond ETF trading leads to excessively underlying corporate bond comovements, a higher corporate bond beta with respect to the super portfolio compared to the aggregate bond market would be expected.

Empirically, I construct a comprehensive panel of corporate bond ETFs and use panel fixed effect regressions with control variables to quantify the impacts of corporate bond ETFs ownership on the intensity of corporate bond return comovement. I find that the higher corporate bond ETFs ownership leads to higher bond return comovement intensity with the aggregate market and with the basket of corporate underlying the same corporate bond ETFs. I control for various bond level characteristics, time fixed effects and bond fixed effects. Specifically, my results indicate that one percentage increase in the corporate bond ETFs ownership leads to an increase of 0.26 in corporate bonds' beta with respect to the aggregate bond portfolio. The positive relationship is pervasive and robust with various specifications and different measurements of corporate bond comovment intensity.

After establishing the positive relationship between bond ETF ownership and corporate bond return comovement, I further investigate if ownership by other traditional institutional investors in corporate bond market has a similar effect on bond return comovement. I compare the effect of corporate bond ETFs ownership with corporate bond ownership by corporate bond mutual fund and insurance companies. It turns out that corporate bond ETFs ownership uniquely leads to increase in bond return comovement. The effect of bond ownership by active bond mutual fund ownership is negative with statistical significance. It is reasonable as active bond mutual funds tend to trade on private information which tends to improve the price efficiency as shown in the first two chapters. Furthermore, corporate bond ownership by insurance companies is insignificant, which is also consistent with institutional backgrounds since insurance companies are majority buy and hold investors.

After establishing a robust relationship between corporate bond ownership by corporate bond ETFs and the underlying bond return comovement, I find that the improved bond return comovment is driven by creation and redemption arbitrage activities of corporate bond ETFs. Demand for bond ETFs results in price pressure that causes the price of ETF shares to decouple from the value of the underlying baskets of corporate bond. Authorized participants of bond ETFs will step in to arbitrage between the ETFs shares and the underlying corporate bond markets to ensure the value of ETFs aligns with the underlying corporate bond value. When corporate bond ETFs are traded at the discounted with the value of ETF shares lower than the value of the underlying corporate bond, authorized participants would purchase the ETFs shares and short the underlying corporate bonds. Similarly, when corporate bond ETFs are traded at the premium with the price of ETF shares higher the value of corporate bonds, they will sell ETF shares and purchase the underlying caproate bond. This arbitrage behavior explains bond return comovement resulting from the corporate bond ownership by corporate bond ETFs. By using the standard deviation of outstanding shares on corporate bond ETFs as a proxy for the intensity of creation and redemption activities, I find that higher creation and redemption activities explain the increase in corporate bond return comvoment.

The rest of the chapter is organized as follows. Section 3.2 presents related literature and contribution. Section 3.3 provided institutional background. Section 3.4 shows main hypotheses development. Section 3.5 presents sample data and variables construction.

Section 3.6 discusses the empirical analysis of impact of corporate bond ETF ownership on the underlying bond return movement. Section 3.7 concludes.

3.2 Contribution to Literature

This chapter contributes to a growing literature on the exchange-traded fund by providing the first evidence about the impact of corporate bond ETFs on bond return comovement. Current studies mostly focus on equity ETFs (Madhavan (2016), Lettau and Madhavan (2018) and Ben-David, Franzoni, and Moussawi (2017)). Israeli, Lee, and Sridharan (2017) and Glosten, Nallareddy, and Zou (2016) study the relationship between ETFs holding and the liquidity of underlying stocks. In addition, studies find that stocks held by equity ETFs increase volatility (Ben-David, Franzoni, and Moussawi (2018)) and return fcomovement (Da and Shive (2018). In the field of corporate bond ETFs, Dannhauser (2017) shows that higher ETF ownership decreases corporate bond yield and Pan and Zeng (2019) studies the liquidity mismatch between bond ETFs and underlying corporate bonds. They draw different conclusions regarding the impact of bond ETFs on bond liquidity. This paper fills the gap in the literature by studying the impacts of corporate bond ETFs on comovement of underlying corporate bonds.

This paper also adds to the large literature on return comovement. Prior studies show that index addition (Barberis, Shleifer, and Wurgler (2005) and Wurgler (2010)) and stock spilt (Green and Hwang (2009)) lead to excess stock return comovement. Peng and Xiong (2006) find that investor attention and categorical learning cause stocks to comoves more with other stocks with similar characteristics. Moreover, correlated trading and ownership behaviors by institutional investors are also found to contribute to excessive stock return comovement (Greenwood and Thesmar (2011), Antón and Polk (2014), Bartram, Griffin, Lim and Ng (2015) and Da and Shive (2018)). Additional studies focus on equity-bond market comovement (Baker and Wurgler (2012), Bao and Hou (2013), and David and Veronesi (2016)). This paper contributes to literature of return comovement by presenting the new evidence on the corporate bond return comovement driven by bond ETF trading and arbitrages behaviors.

This paper is close to Da and Shive (2018) which shows that equity ETFs ownership contributes to stocks comovement. My paper differs from Da and Shive (2018) by focusing on corporate bond ETFs and corporate bond return comovement. It shows how the return comovement relates to corporate bond ETF arbitrage activities across the markets. Different form equity ETFs, corporate bond ETFs are featured with liquidity mismatch. Both equity ETFs and the underlying stocks are traded on the exchange. In contrast, corporate bond ETFs is traded as a stock on the exchange with its underlying trading over the counter in a relatively less liquid environment. Da and Shive (2018) do not find stock comovement relates strongly to equity ETFs creation and redemption arbitrage activities. However, due to the differences between Equity ETFs and Corporate bond ETFs, I find that corporate bond return comovement is driven significantly by bond ETF creation and redemption arbitrage activities.

3.3 Corporate bond ETF Institutional Background

3.3.1 Corporate Bond ETF

Corporate bond exchange-traded funds are marketable securities that track indexes or a basket of fixed income assets. Corporate bond ETFs are managed passively, as weights of bonds held by bond ETFs are pre-determined according to rules or the indexes. Moreover, they are a hybrid of closed-end mutual funds and open-end mutual funds. On one hand, similar to open-end mutual funds, they are bought and sold at their net asset value of the underlying assets at the end of each trading day. On the other hand, similar to closed-end funds, they are traded within the trading day at the market-determined price which may be different from their net asset value. In addition, bond ETF shares are allowed to be created and redeemed at the end of the trading day for the funds' net asset value. In contrast to open-end funds, the creation and redemption activities occur only between bond ETF issuers and authorized anticipants, who are primarily broker-dealers. This is the unique feature of ETFs, which is also referred to as the primary ETF market.

The process to establish an exchange-traded fund is as follows. ETF issuers called as the fund sponsors will first specify the investment objective, indexes to be tracked and the tracking methodology of funds. They would then reach out to authorized participants to either borrow or purchase the relevant underlying securities for them, in order to mimic the composition of published indexes. Authorized participants are large broker and dealer in the market with access to various underlying securities. Authorized participants hand the underlying securities to ETF sponsors and in return, ETF sponsors hand creation units to Authorized participants. One creation unit often equals to 50000 shares. Authorized participants would then sell the create unit to the stock market, as a normal stock. This is referred as the secondary market, where investors purchase ETF shares.

3.3.2 Bond ETF Creation and Redemption Arbitrage Activities

The bond ETF creation/redemption process is uniquely designed to ensure that the price of ETF shares equals the net asset value of the underlying corporate bonds held by the bond ETF. Authorized participants start to assume the role of arbitraguer across the markets when a bond ETF is established. If ETF shares are traded above the net asset value of the underlying assets, authorized participants will short sell the ETF shares and purchase the underlying assets. It makes the ETF price align with the value of underlying securities. At the ending of the trading day, authorized participants offset their short position by handing the purchased assets to ETF sponsors in exchange for ETF shares. Similarly, if ETF shares are traded at a discount with the price of ETF shares below the net asset value of the underlying assets, authorized participants will sell the underlying assets they hold and purchase more ETF shares instead in order to ensure two prices close to each other. At the end of the day, they turn in the extra shares they purchase during the day to ETF sponsors in exchange for underlying assets.

3.4 Main Hypothesis Development

In this section, I present the following three hypotheses based on the institutional details of corporate bond ETFs described in the previous section.

Hypothesis I: Corporate bond ETF ownership increases underlying corporate bond return comovement.

Corporate bond ETFs trade on the exchange like a stock, with their underlying corporate bonds trading over the counter. Demand for bond ETF shares might migrate to the underlying bonds. When demand for ETFs results in price pressure which causes ETF prices to decouple from the fundamental values, arbitrage activities will take the opposite position between bond ETFs and the underlying corporate bonds. Bond ETF helps transmit the non-fundamental shocks to the underlying bonds. As a result, the bonds owned by ETFs comove with each other more than what they are supposed to, based on their common exposure to the fundamental information.

Hypothesis II: Enhanced bond return comovement is driven by corporate bond ETF creation and redemption arbitrage activities.

Hypothesis III: Corporate bond ownership by traditional institutional investors like bond mutual funds and insurance companies do not increase the underlying corporate bond return comovement.

Corporate bond ownership by bond mutual funds is not expected to increase bond return comovement since bond mutual funds trade actively on private information. Their trading activities help improve bond pricing efficiency as shown in the previous chapters. Moreover, corporate bond ownership by insurance companies is also not expected to increase bond return comovement because insurance companies are primarily buy and hold investors.

3.5 Data Construction

My main data on corporate bond exchange-traded funds is from the following databases: Enhanced Historic Trade Reporting and Compliance Engine (TRACE Enhance) and Fixed Investment Securities Database(FISD) for bond level characteristics, Thomson Reuters Lipper eMaxx, Center for Research in Security Price (CRSP) Survivorship Bias-Free Mutual Fund Database and Bloomberg for corporate bond exchanged trade fund level information and Datastream for pricing factors.

3.5.1 Corporate bond pricing and trading data

The comprehensive corporate bond transaction data comes from the TRACE Enhance database, which is introduced in 2002 covering detailed high frequency transactional level data in the corporate bond market. Though the first ETF was introduced in 2002 around the same as that of the introduction of trace, the assets managed by ETFs were trivial in the first decade. In order to analyze the effects of ETF during a period when asset held by ETFs are significant, the sample starts in 2011. For all the observable security level eight-digit CUSIP identified in TRACE database, I match them with FISD to obtain detailed bond issue and bond issuer level characteristics. Following the tradition in the literature (Dannhauser (2017), Choi and Kronlund (2017), and Cai, Song, Dan and Yi (2019)), I exclude all perpetual, preferred, convertible, asset backed, variable rate, sinking fund, and Rule 144A securities. I also apply the methodology in Dick-Nielsen

(2014) to filter all the erroneous trade and correct for double counted transactions and liquidity biases problem in the transactional database. Moreover, I adopt the highest rating of major credit rating agencies Moody's, S&P, and Fitch as the final bond rating for each security. The advantage of high frequency transaction data allows me to calculate various liquidity and bond return measurement. CRSP Bond return is also used as an additional return source.

3.5.2 Corporate bond exchanged fund holding and trading related data

Corporate bond ETF is identified using both CRSP Mutual fund database and Thomson Reuters eMaxx database. Mutual funds are required to report their quarter holding information. CRSP Mutual fund database gathers mutual fund prospectuses and fund fact sheets. The ticker and name of corporate bond exchange-traded funds are acquired using et_flag and crsp_obj_cd filters in CRSP. I also supplement it with the list from ETF website. After generating the complete list of corporate bond exchange traded ETF, I merge the fund list with the eMaxx database to retrieve the quarterly holding. The eMaxx database contains the comprehensive corporate bond trading and holding positions by institutional investors including corporate bond exchange-traded fund, mutual funds, pension funds and insurance company on a quarterly basis. Prior literature (Massa and Zhang (2012), Becker and Ivashina (2015), Choi and Kronlund (2017) and Cai, Song, Dan and Yi (2017)) also use eMaxx for corporate bond market studies as eMaxx is a high quality database which includes all defunct funds and free of survivorship bias and reporting biases. This paper is interested in the holding of corporate bond exchanged traded. I also compare the effect of holding by corporate bond ETF with those of other institutional investors. In doing, I create a detailed bond level panel data for all the corporate bond held by corporate bond ETFs.

3.5.3 Corporate bond exchanged trade fund level characteristics

As mentioned previously, I apply the CRSP Lipper Object code to filter out all the mutual fund in eMaxx in order to identify the primary corporate bond exchange-traded funds. The fund level information is extracted from CRSP mutual fund database to gather fund level characteristics. Unfortunately, eMaxx and CRSP don't share the similar fund identifier code. Fund level merge linkage is done by matching fund legal names. I first adopt the searching algorithm and fuzzy matching algorithm from computer science literature to complete the first-round matching and I manually check each name to guarantee the accuracy. Both steps result in 125 unique corporate bond exchange-traded funds.

After generating the accurate fund level link, I retrieve exhaustive Bond ETF fund level information which includes mutual fund returns, flow, investment style, total assets, and cash holdings. Holding the same portfolio, fund share classes from the same fund only differs by fee structures and expense ratios, I further group all variables from CRSP share classes to fund level using value weighted method. The total net asset in the previous period is served as weight. Additionally, I also use the Bloomberg terminal to obtain daily and monthly fund level information about fund trading the arbitrage activities. It includes total outstanding shares, turnover, daily price, trading volume, and etc.

3.5.4 Market return and bond mutual fund benchmark index return

From DataStream, I acquire several mutual fund benchmarks which include Vanguard Total Bond Market Index Fund return (Ng, Wang, and Wang (2015)). In addition, I obtain aggregate market returns from CRSP Fama French Factors Library. The final sample includes 65413 bond issue level quarterly observations for bond issue level study. It includes 7210 unique bonds held by corporate bond exchange-traded funds.

3.6 Empirical Results

In this section, I present the empirical analysis regarding the relationship between corporate bond ETF ownership and corporate bond return comovement. I first begin with the panel fixed effect regression to model this relationship, which shows that corporate bond ETF ownership leads to higher bond return comovement. Moreover, I compare the impacts on bond return comovement by bond ETF ownership with corporation bond ownership of other traditional institutional investors in the corporate bond markets including bond mutual funds and insurance companies. It turns out that bond ownership by corporate bond ETFs uniquely contributes to underlying bond return comovement, in contrast to the negative or insignificant impacts of mutual fund and pension fund ownership on bond return comovement. I conclude this section by studying the mechanism that could possibly drive this positive relationship between corporate bond ETF ownership and bond return co-movement from the perspective of bond ETF creation and redemption arbitrage activities.

3.6.1 Positive impacts of bond ownership by corporate bond ETFs on bond return comovement

Utilizing the panel fixed effect model, I examine the effect of corporate bond ETF ownership on the underlying corporate bond return comovement:

$$ReturnComovement_{i,t+1} = \beta_1 ETF\ Onwership_{i,t} + Controls_{i,t} + FixedEffect_{i,t} + \epsilon_{i,t}$$

where $ReturnComovement_{i,t+1}$ is corporate bond return comovement with the aggregate bond portfolio, which is the coefficient of the corporate bond's monthly excess returns on monthly bond market excess returns or on the monthly aggregate portfolio of ETFs that own the underlying security. The variable is constructed based on the return comovement measure commonly adopted in the literature (Da and Shive (2018)). $ETF\ Onwership_{i,t}$ is the proportion of the underlying corporate bonds owned by corporate bond ETFs. I include various corporate bond level characteristic as control variables. For example, I control for bond rating, time to maturity and turnover which might potentially influence the bond return comovement with the aggregate bond portfolio. I also include bond fixed effects and time fixed effect to further control for other factors that could correlate with bond return comovement.

[Insert Table 3.1 here]

The variable of interest is β_1 which measures the sensitivity between corporate bond ETF ownership and caproate bond return co-movement. The positive β_1 indicates that corporate bond ETF ownership improves bond return co-movement. Table 3.1 presents the empirical estimation results. Bond return comovement is measured in two different specifications. The first two columns in Table 3.1 report the effect on bond return co-movement with the aggregate bond market and last two columns in Table 3.1 report the effect on bond return comovement with aggregate ETF portfolio that owns the underlying bond. I find that bond ownership by corporate bond funds leads to the increase in corporate bond return comovement in all the specifications. More specifically, one percent increase in corporate bond ETF increases the bond return co-movement measured by beta on the aggregate bond market by 0.036 and 0.026 if fixed effects are included. One percent increase in corporate bond ETF increases the bond return comovement measured by beta on ETF portfolio by 0.012 and 0.009 respectively. It varies from 0.090 to 0.036 of increase on the beta of the corporate bond return with respect to the aggregate bond portfolio. Moreover, the impacts of corporate bond ETF on bond return comovement are statistically significant and economically significant. From the perspective of portfolio managers, 0.009 increase on beta on ETF portfolio indicates that portfolio managers need to enhance their portfolio return by 90 basis points so as to generate the same alpha. Therefore, corporate bond ETF ownership has a substantial positive impact on the corporate bond return comovement, which lends support to my hypothesis.

3.6.2 Impacts of bond ownership by other traditional institutional investors on bond return comovement

In this section, I want to test if bond ownership by corporate bond ETFs uniquely contribute the positive impact on corporate bond return results from corporate bond ETF ownership uniquely. I compare the effects of bond ownership by mutual funds and insurance companies with bond ownership by corporate bond ETF ownership. I run the following panel fixed effects regressions:

$$ReturnComovement_{i,t+1} = \beta_1 ETF\ Onwership_{i,t} + \beta_2 MutualFundOnwership_{i,t} + \beta_3 InsuranceFundOnwership_{i,t} + Controls_{i,t} + FixedEffect_i + \epsilon_{i,t}$$

where $ReturnComovement_{i,t+1}$ is corporate bond return comovement. I mainly focus on the bond return co-movement with aggregate bond portfolio, which is the beta coefficient of corporate bond's excess returns with respect to aggregate bond portfolio excess returns. $ETF\ Onwership_{i,t}$ is the proportion of corporate bonds owned by corporate bond ETFs. $MutualFundOnwership_{i,t}$ is the proportion of the underlying corporate bonds held by corporate bond mutual funds. $InsuranceFundOnwership_{i,t}$ is the proportion of the underlying corporate bonds owned by insurance companies. Various corporate bond level characteristics are included as control variables. I control for bond rating, time to maturity and turnover. In addition, I include bond fixed effects and time fixed effect to further control for factors that potentially influence corporate bond return comovement. I intend to compare the influence of corporate bond ownership by various institutional investors on corporate bond return co-movement.

[Insert Table3.2 here]

I am interested in the coefficients of $\beta_1, \beta_2,$ and $\beta_3$which measures the influence of corporate bond institutional investors ownership on their underlying caproate bond return comovement. Table 3.2 presents the empirical results. In contrast to the positive impacts by corporate bond ETF ownership on bond return comovement, corporate bond ownership by other institutional investors results in either negative impacts or no impacts. Specifically, one percent increase in corporate bond ownership by mutual funds decreases corporate bond return comovement by 5 basis points. Corporate bond ownership by insurance companies shows no significant effects on corporate bond return comovement. It is consistent with the institutional background that insurance companies are primary buy-and-hold investors in the corporate bond market and they do not trade frequently compared to other participants in the market. Thus, their corporate bond ownership bond does not show significant impacts on the corporate bond market. Table 3.2 indicates that comparing with other institutional investors in corporate bond market, corporate bonds ownership by corporate bond ETFs is the unique one that contributes to corporate bond return comovement. This further confirms the substantial and important economic impacts of corporate bond ETFs on the underlying corporate bonds.

3.6.3 Positive impacts of corporate bond ETF ownership on their underlying bond return comovement are driven by bond ETF arbitrage activities.

After showing that corporate bond ETFs ownership leads to increased bond return comovement, I want to examine the potential mechanism that drives this positive relationship. Specifically, I intend to test if the bond return comovement associated with corporate bond ETF ownership is due to the creation and redemption arbitrage activities of corporate bond ETFs. When arbitrageurs of corporate bond ETFs conduct arbitrage activities across the markets and trade upon the underlying securities, it thus leads to the returns of related bonds to comove more with each other. ETF arbitrage activities occur through bond ETF creation and redemption. When corporate bond ETFs trades at a discount, authorized participants of bond ETFs will buy bond ETFs shares and short sell the underlying corporate bonds. At the end of each trading day, authorized participants return the corporate bond shares they purchase to the ETF sponsors in order to offset their short positions. Similarly, when corporate bond ETFs trades at a premium, authorized participants of bond ETFs will buy the underlying assets and short sell the corporate bond ETFs shares. They hand in ETFs they purchase to bond ETF sponsors in exchange for corporate bonds to offset their position. During the creation and redemption process, the bonds owned by corporate bond ETFs are traded together mechanically. It hence leads to higher bond return comovement. I run the following fixed effects regression:

$$ReturnComovement_{i,t+1} = \beta_1 ETF\ Creation\&Redemption_{i,t} + Controls_{i,t} +$$

$$FixedEffect_{i,t} \ + \epsilon_{i,t}$$

where $ReturnComovement_{i,t+1}$ is corporate bond return comovement. The bond return comovement with the aggregate bond portfolio is measured by the beta of the corporate bond's excess returns with respect to the aggregate bond portfolio. $ETF\ Creation\&Redemption_{i,t}$ captures the creation and redemption activities of corporate bond ETFs. It is measured by the standard deviation of corporate bond ETFs daily outstanding shares following Da and Shive (2018). Various corporate bond level characteristics like bond rating, time to maturity and turnover are used as control variables. In addition, I also include bond fixed effects and time fixed effect to more rigorously control for factors that could bias the result by correlating with corporate bond return comovement. I intend to examine if corporate bond return comovement driven by ETF ownership results from bond ETF creation and redemption arbitrage activities.

[Insert Table 3.3 here]

I am interested in $\beta_1$which examines if the positive relationship between corporate bond ETF and corporate bond return comovement is driven by corporate bond ETF creation and redemption activities. Table 3.3 reports that consistent with my hypothesis, ETF creation and redemption arbitrage activities serve as the mechanism behind this positive relationship. One standard deviation increase in the ETF creation and redemption arbitrage activities associates with intensified corporate bond return comovement by 138 basis points. This suggests that trading and arbitrage activities done by corporate bond

ETFs plays an influential role in the corporate bond markets. It drives the return of related corporate bond to comove more with each other. This lends support to my hypothesis in the prior section. Thus, bond ownership by corporate bond ETFs increases their underlying bond return comovement via bond ETFs' creation and redemption arbitrage activities.

3.7 Conclusion

This paper investigates the impacts of corporate bond ETFs on the price movement of their underlying corporate bonds. The rapid growth of corporate ETFs offers investors diversified and low-cost investment options. It also makes the corporate bond index more accessible and cheaper than ever before. However, ETF trading and arbitrage activities have potential to transmits the non-fundamental shock to their underlying corporate bond securities. I find that corporate bond ownership by corporate bond ETFs leads to high corporate bond return comovement. Moreover, by comparing the corporate bond ETF ownership with bond ownership by other traditional investors in the corporate bond market, I show that corporate bond ETFs uniquely contributes to the underlying bond return comovement. In contrast, bond ownership by corporate bond mutual has a negative impact on their underlying return comovement. Bond ownership by insurance companies has an insignificant impact on the bond return comovement as they are primarily buy-and-hold investors in the corporate bond market. Furthermore, I examine the creation and redemption arbitrages activities of corporate bond ETFs and I show that bond return comovement due to bond ETF ownership is driven by corporate bond ETFs' creation and redemption activities. This paper contributes to the literature by providing the first empirical evidence of corporate bond return comovement driven by bond ETF trading activities.

3.8 Figures and Tables

Table 3.1: Relationship between corporate bond return comovement and corporate bond ETF ownership

This table presents the relationship between corporate bond ETF ownership and corporate bond return comovement. The table reports both bond return comovement with aggregate bond market portfolio and bond return comovement with aggregate bond ETF portfolio. Bond return comovement measurement is constructed following Da and Shive (2018). Bond return comovement with the aggregate bond market portfolio is reported in the first two columns of the table. It is measured by the beta of the corporate bond's excess return with respect to the bond market excess return. Column three and column four report bond return comovement with aggregate ETF portfolio which owns the underlying securities. Bond ETF ownership is the proportion of corporate bond owned by corporate bond ETFs. Bond Rating is the max of ratings coming from three credit rating agencies: Moody's, S&P and Fitch. Time to Maturity is time to maturity. Turnover is the average daily turnover of the bond over the month. Bond level fixed effects and time fixed effects are included in the regression. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Bond Return Co-movement			
	Beta on Aggregate Bond Portfolio	Beta on Aggregate Bond Portfolio	Beta on Aggregate Bond ETF Portfolio	Beta on Aggregate Bond ETF Portfolio
	(1)	(2)	(3)	(4)
ETF Ownership	0.036***	0.026**	0.012***	0.009***
	(0.012)	(0.012)	(0.0004)	(0.0004)
Bond Rating	0.106***	0.091***	0.007***	0.003***
	(0.022)	(0.022)	(0.001)	(0.001)
Time to Maturity	0.012***	0.001	0.001***	-0.001
	(0.001)	(0.003)	(0.00003)	(0.002)
Turnover	-0.031*	-0.026	-0.046***	-0.038***
	(0.017)	(0.017)	(0.001)	(0.001)
Bond Fixed Effect	Yes	Yes	Yes	Yes
Time Fixed Effect	No	Yes	No	Yes
Observations	59,266	59,266	59,266	59,266
R^2	0.005	0.008	0.101	0.070

Table 3.2: Relationship between corporate bond return comovement and bond ownership by other institutional investors including mutual funds and insurance company

This table shows the comparison of the impacts on corporate bond return by various institutional investors including corporate bond ETF, mutual fund and insurance companies. Bond return comovement measure is constructed following Da and Shive (2018). Bond return comovement with the aggregate bond portfolio is measured by the beta of the corporate bond's excess return with respect to the aggregate bond market excess return. Bond ETF Ownership is the proportion of corporate bond owned by corporate bond ETFs. Mutual Fund Ownership is the proportion of corporate bond owned by mutual funds. Insurance Fund Ownership is the proportion of corporate bond owned by insurance companies. Bond Rating is the max of ratings coming from three credit rating agencies: Moody's, S&P and Fitch. Time to Maturity is time to maturity. Turnover is the average daily turnover of the bond over the month. Bond level fixed effects and time fixed effects are included in the regression. Stars represent statistical significance. (***p<0.001, **p<0.005, *p<0.1, respectively).

	Bond Return Co-movement with Aggregate Bond Portfolio			
	(1)	(2)	(3)	(4)
ETF Ownership	0.036***	0.026**		
	(0.012)	(0.012)		
Mutual Fund Ownership			-0.005*	
			(0.003)	
Insurance Fund Ownership				-0.004
				(0.003)
Bond Rating	0.106***	0.091***	0.090***	0.084***
	(0.022)	(0.022)	(0.022)	(0.022)
Time to Maturity	0.012***	0.001	0.001	0.001
	(0.001)	(0.003)	(0.003)	(0.003)
Turnover	-0.031*	-0.026	-0.011	-0.010
	(0.017)	(0.017)	(0.015)	(0.015)
Bond Fixed Effect	Yes	Yes	Yes	Yes
Time Fixed Effect	No	Yes	Yes	Yes
Observations	59,266	59,266	59,266	59,266
R^2	0.005	0.008	0.008	0.008

Table 3.3: Corporate Bond Return Comovement and Bond ETF Creation and Redemption Activities

This table shows corporate bond return comovement driven by corporate bond ETF creation and redemption arbitrage activites. Bond return co-movement is measured by the coefficient of the bond's excess returns on aggregate bond market excess return in the same month. Bond ETF Ownership is the proportion of corporate bond owned by corporate bond ETFs. Mutual Fund Ownership is the proportion of corporate bond owned by mutual funds. Insurance Fund Ownership is the proportion of corporate bond owned by insurance companies. Bond Rating is the max of ratings coming from three credit rating agencies: Moody's, S&P and Fitch. Time to Maturity is time to maturity. Turnover is the average daily turnover of the bond over the month. Bond level fixed effects and time fixed effects are included in the regression. Stars represent statistical significance. (****p*<0.001, ***p*<0.005, **p*<0.1, respectively)

	Return Co-movement
	(1)
ETF Creation/Redemption	0.138***
	(0.032)
Bond Rating	0.089***
	(0.022)
Time to Maturity	0.002
	(0.003)
Turnover	-0.065***
	(0.020)
Bond Fixed Effect	Yes
Time Fixed Effect	Yes
Observations	59,266
R^2	0.008

Figure 3.1: Total Net Asset Value of Bond Exchange-Traded Funds

This figure presents the total net assets of bond exchange-traded funds in the U.S. The data source is from Investment Company Institution Fact Book 2019.

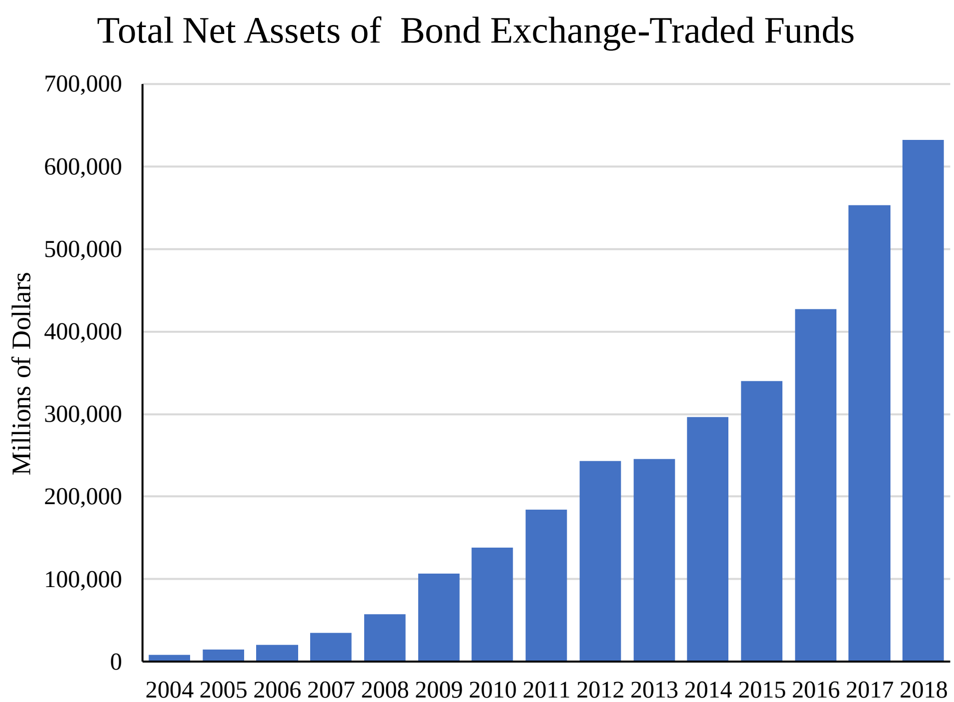
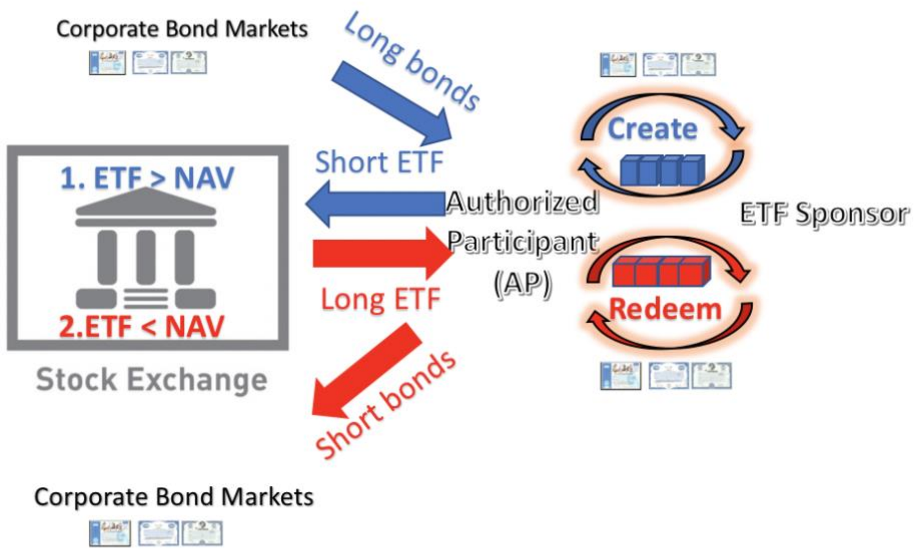


Figure 3.2: Corporate Bond ETF Creation and Redemption Arbitrage Activities

This figure presents the Corporate Bond ETFs' creation and redemption process. Authorized participants arbitrage across the markets between corporate bond ETF shares and the underlying corporate bonds. The arbitrage activities done by authorized participants is to ensure that the price of ETF shares equal to the value of underlying corporate bonds.



Bibliography

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, 2012, The Network Origins of Aggregate Fluctuations, *Econometrica* 80, 1977–2016.

Adrian, Tobias, Michael J. Fleming, Or Shachar, and Erik Vogt, 2016, Market Liquidity after the Financial Crisis. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Ahern, Kenneth R., 2016, Information Networks: Evidence from Illegal Insider Trading Tips. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Amihud, Yakov, 2002, Illiquidity and Stock Returns: Cross-section and Time-series Effects, *Journal of Financial Markets*, 31–56.

Antón, Miguel, and Christopher Polk, 2014, Connected Stocks, *The Journal of Finance* 69, 1099–1127.

Babus, Ana, and Péter Kondor, 2018, Trading and Information Diffusion in Over-the-Counter Markets, *Econometrica* 86, 1727–1769.

Baker, Malcolm, and Jeffrey Wurgler, 2012, Comovement and Predictability Relationships Between Bonds and the Cross-section of Stocks, *The Review of Asset Pricing Studies* 2, 57–87.

Bao, Jack, and Kewei Hou, 2013, Comovement of Corporate Bonds and Equities. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which Factors Matter to Investors? Evidence from Mutual Fund Flows, *The Review of Financial Studies* 29, 2600–2642.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283–317.

Bartram, Söhnke M., John M. Griffin, Tae-Hoon Lim, and David T. Ng, 2015, How Important Are Foreign Ownership Linkages for International Stock Returns?, *Review of Financial Studies* 28, 3036–3072.

Becker, Bo, and Victoria Ivashina, 2015, Reaching for Yield in the Bond Market, *The Journal of Finance* 70, 1863–1902.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2017, Exchange-Traded Funds, *Annual Review of Financial Economics* 9, 169–189.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2018, Do ETFs Increase Volatility?, *The Journal of Finance* 73, 2471–2535.

Berk, Jonathan B., and Jules H. van Binsbergen, 2016, Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119, 1–23.

Blake, Christopher R., Edwin J. Elton, and Martin J. Gruber, 1993, The Performance of Bond Mutual Funds, *The Journal of Business* 66, 371–403.

Bonacich, Phillip, 2007, Some unique properties of eigenvector centrality, *Social Networks* 29, 555–564.

Booth, G. Geoffrey, Umit G. Gurun, and Harold Zhang, 2014, Financial networks and trading in bond markets, *Journal of Financial Markets* 18, 126–157.

Brown, Nerissa C., Kelsey D. Wei, and Russ Wermers, 2013, Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices, *Management Science* 60, 1–20.

Cai, Fang, Song Han, Dan Li, and Yi Li, 2017, Institutional Herding and Its Price Impact: Evidence from the Corporate Bond Market. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Cai, Fang, Song Han, Dan Li, and Yi Li, 2019, Institutional herding and its price impact: Evidence from the corporate bond market, *Journal of Financial Economics* 131, 139–167.

Cai, Fang, Song Han, Dan Li, and Yi Li, 2019, Institutional herding and its price impact: Evidence from the corporate bond market, *Journal of Financial Economics* 131, 139–167.

Chen, Qi, Itay Goldstein, and Wei Jiang, 2007, Price Informativeness and Investment Sensitivity to Stock Price, *Review of Financial Studies* 20, 619–650.

Chen, Qi, Itay Goldstein, and Wei Jiang, 2010, Payoff complementarities and financial fragility: Evidence from mutual fund outflows, *Journal of Financial Economics* 97, 239–262.

Choi, Jaewon, and Yesol Huh, 2017, Customer Liquidity Provision: Implications for Corporate Bond Transaction Costs, *Finance and Economics Discussion Series* 2017.

Choi, Jaewon, and Mathias Kronlund, 2017, Reaching for Yield in Corporate Bond Mutual Funds, *The Review of Financial Studies*.

Cohen, Lauren, Andrea Frazzini, and Christopher Malloy, 2008, The Small World of Investing: Board Connections and Mutual Fund Returns, *Journal of Political Economy* 116, 951–979.

Cutura, Jannic, Gianpaolo Parise, and Andreas Schrimpf, 2017, Do Zombie Funds Gamble for Resurrection? Evidence from the US Corporate Bond Market. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Da, Zhi, and Sophie Shive, 2018, Exchange traded funds and asset return correlations, *European Financial Management* 24, 136–168.

Dannhauser, Caitlin D., 2017, The impact of innovation: Evidence from corporate bond exchange-traded funds (ETFs), *Journal of Financial Economics* 125, 537–560.

David, Alexander, and Pietro Veronesi, 2016, The Economics of the Comovement of Stocks and Bonds, *Handbook of Fixed-Income Securities* (John Wiley & Sons, Ltd).

Di Maggio, Marco, Francesco A. Franzoni, Amir Kermani, and Carlo Sommavilla, 2017, The Relevance of Broker Networks for Information Diffusion in the Stock Market. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song, 2017, The value of trading relations in turbulent times, *Journal of Financial Economics* 124, 266–284.

Di Maggio, Marco, and Alireza Tahbaz-Salehi, 2015, Collateral shortages and intermediation networks. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Dick-Nielsen, Jens, 2014, How to Clean Enhanced TRACE Data. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Dick-Nielsen, Jens, Peter Feldhütter, and David Lando, 2012, Corporate bond liquidity before and after the onset of the subprime crisis, *Journal of Financial Economics* 103, 471–492.

Falato, Antonio, Ali Hortacsu, Dan Li, and Chaehee Shin, 2016, Fire-Sale Spillovers in Debt Markets. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Feldhütter, Peter, 2012, The Same Bond at Different Prices: Identifying Search Frictions and Selling Pressures, *The Review of Financial Studies* 25, 1155–1206.

Glosten, Lawrence R., Suresh Nallareddy, and Yuan Zou, 2016, ETF Activity and Informational Efficiency of Underlying Securities. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Goldstein, Itay, Hao Jiang, and David T. Ng, 2017, Investor flows and fragility in corporate bond funds, *Journal of Financial Economics* 126, 592–613.

Green, T. Clifton, and Byoung-Hyoun Hwang, 2009, Price-based return comovement, *Journal of Financial Economics* 93, 37–50.

Greenwood, Robin, Samuel G. Hanson, and Jeremy C. Stein, 2015, A Comparative-Advantage Approach to Government Debt Maturity, *The Journal of Finance* 70, 1683–1722.

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

Han, Bing, and Liyan Yang, 2013, Social Networks, Information Acquisition, and Asset Prices, *Management Science* 59, 1444–1457.

Harris, Lawrence, 2015, Transaction Costs, Trade Throughs, and Riskless Principal Trading in Corporate Bond Markets. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2004, Social Interaction and Stock-Market Participation, *The Journal of Finance* 59, 137–163.

Israeli, Doron, Charles M. C. Lee, and Suhas A. Sridharan, 2017, Is there a dark side to exchange traded funds? An information perspective, *Review of Accounting Studies* 22, 1048–1083.

Ivković, Zoran, and Scott Weisbenner, 2007, Information Diffusion Effects in Individual Investors' Common Stock Purchases: Covet Thy Neighbors' Investment Choices, *The Review of Financial Studies* 20, 1327–1357.

Jackson, Matthew O., 2009, An Overview of Social Networks and Economic Applications∗, .

Korajczyk, Robert, and Ronnie Sadka, 2008, Pricing the commonality across alternative measures of liquidity, *Journal of Financial Economics* 87, 45–72.

Lesmond, David A., Joseph P. Ogden, and Charles Trzcinka, 1999, A New Estimate of Transaction Costs, *Review of Financial Studies* 12, 1113–41.

Lettau, Martin, and Ananth Madhavan, 2018, Exchange-Traded Funds 101 for Economists, *Journal of Economic Perspectives* 32, 135–154.

Madhavan, Ananth N., 2016, *Exchange-Traded Funds and the New Dynamics of Investing* (Oxford University Press).

Massa, Massimo, and Lei Zhang, 2012, CDS and the Liquidity Provision in the Bond Market. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Ng, Lilian, Crystal X Wang, and Qinghai Wang, 2015, Risk Taking and Performance of Bond Mutual Funds, .

O'Hara, Maureen, Yihui Wang, and Xing (Alex) Zhou, 2016, The Execution Quality of Corporate Bonds. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Ozsoylev, Han N., and Johan Walden, 2011, Asset pricing in large information networks, *Journal of Economic Theory* 146, 2252–2280.

Ozsoylev, Han N., Johan Walden, M. Deniz Yavuz, and Recep Bildik, 2014, Investor Networks in the Stock Market, *The Review of Financial Studies* 27, 1323–1366.

Pan, Kevin, and Yao Zeng, 2019, ETF Arbitrage Under Liquidity Mismatch. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Pareek, Ankur, 2012, Information networks: Implications for mutual fund trading behavior and stock returns. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics*, 40.

Rossi, Alberto, David P. Blake, Allan Timmermann, Ian Tonks, and Russ Wermers, 2015, Network Centrality and Delegated Investment Performance. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Sias, Richard W., 2004, Institutional Herding, *Review of Financial Studies* 17, 165–206.

Singh, Vivek, 2013, Did institutions herd during the internet bubble?, *Review of Quantitative Finance and Accounting* 41, 513–534.

Walden, Johan, 2014, Trading, Profits, and Volatility in a Dynamic Information Network Model, *SSRN Electronic Journal*.

Wurgler, Jeffrey, 2010, On the Economic Consequences of Index-Linked Investing. Working Paper, National Bureau of Economic Research.

Zhu, Yuyuan, 2016, Social Connections and Information Production: Evidence from Mutual Fund Portfolios and Performance, *SSRN Electronic Journal*.